



PA (HARD ) HARVARD COLLEGE.  
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.  
XX  
XX Ingham PW, McMahon AP, Tablin CJ;  
PI  
XX WPI: 1995-255060/33.  
DR N-PSDB: AAO91636.  
XX  
XX Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful  
PT to treat degenerative nervous system disorder(s) and in gene  
PT therapy.  
XX  
PS Claim 17; Page 133-35; 210pp; English.  
XX  
XX The sequence represents a chicken sonic hedgehog protein,  
CC homologous to a Drosophila hedgehog protein (AAR77337), and is  
CC encoded by a cDNA isolated from a stage 22/22 limb bud cDNA  
CC library. Probes and primers derived from the sonic hedgehog gene  
CC may be used as diagnostic agents for neuromuscular, autonomic or  
CC central nervous system disorders, and the gene may also be used in  
CC gene therapy. Antibodies generated from the protein may be used  
CC as therapeutic or research reagents.  
XX  
XX  
SQ Sequence 425 AA;  
Query Match 100.0%; Score 2218; DB 16; Length 425;  
Best Local Similarity 100.0%; Pred. No. 1.7e-227;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WEMLLRILLYGFCIALVSSGLTGPGRIGRRPKKLTPLAYQFIPNVAEKLIG 60  
DB 1 WEMLLRILLYGFCIALVSSGLTGPGRIGRRPKKLTPLAYQFIPNVAEKLIG 60  
OY 61 ASGREGKITRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOM 120  
DB 61 ASGREGKITRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOM 120  
OY 121 PCVKLRVTEGMEDEGHSEESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYYES 180  
DB 121 PCVKLRVTEGMEDEGHSEESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYYES 180  
OY 181 KAHHCISVKAENSVAAKSGCGPFSATVHLEHGRTKIVKDISPGDRVLAADADGRILYSD 240  
DB 181 KAHHCISVKAENSVAAKSGCGPFSATVHLEHGRTKIVKDISPGDRVLAADADGRILYSD 240  
OY 241 FLTFLDRMDSRKLFYIETROPARALLTAHLLFVAPOHNSBATGSGOALFASNV 300  
DB 241 FLTFLDRMDSRKLFYIETROPARALLTAHLLFVAPOHNSBATGSGOALFASNV 300  
OY 301 KPGQAVYVLGEGGQULLPASVHVSGLREBASAVAPLTAOGTILINRYLASCYAVIEEHS 360  
DB 301 KPGQAVYVLGEGGQULLPASVHVSGLREBASAVAPLTAOGTILINRYLASCYAVIEEHS 360  
OY 361 WAHMAFAPRLAQLLALCPDGAIPTAATTTGTHWYSRLYRIGSWLQSDALHPLGM 420  
DB 361 WAHMAFAPRLAQLLALCPDGAIPTAATTTGTHWYSRLYRIGSWLQSDALHPLGM 420  
OY 421 VAPAS 425  
DB 421 VAPAS 425  
RESULT 2  
ID AAY05854 standard; Protein: 425 AA.  
XX  
XX AAY05854;  
AC  
XX  
XX 02-AUG-1999 (first entry)  
DT  
XX Chicken Sonic hedgehog Shh protein.  
DE  
XX Sonic hedgehog; Shh; chicken; epithelial tissue; epithelium;  
KM

KM cutaneous tissue; skin; hair; wound healing; vulneryary;  
KM burn; skin grafting; pressure sore; ulcer; ulcerative colitis;  
KM alopecia; psoriasis; keratosis; acne; comedogenic lesion;  
KM folliculitis; pseudofolliculitis; keratoacanthoma; callositles;  
KM Darier's disease; scar; autoimmune disease; pemphigus;  
KM epidermolysis; lupus lesion; desquamative lesion; carcinoma;  
KM therapy; hedgehog therapeutic; plc therapeutic; patched.  
XX  
XX Gallus domesticus.  
XX  
XX WO9920298-A1.  
XX  
XX 29-APR-1999.  
XX  
XX 20-OCT-1998; 98WO-US22227.  
XX  
XX 11-SEP-1998; 98US-0151999.  
XX 20-OCT-1997; 97US-0955552.  
XX  
XX (OMTO-) ONTOGENY INC.  
PA  
XX  
XX Wang EA;  
PI  
XX WPI: 1999-288170/24.  
DR N-PSDB: AAX25617.  
XX  
XX Use of hedgehog polypeptides on patched therapeutics  
PT  
XX  
XX Claim 26; Page 119-120; 146pp; English.  
PS  
XX  
XX The present sequence represents chicken Sonic hedgehog protein  
CC Shh. The invention relates to a method for modulating the growth  
CC state an epithelial cell by ectopically contacting the epithelial  
CC cell, in vitro or in vivo, with a hedgehog therapeutic (i.e. a  
CC hedgehog polypeptide or gene therapy construct) or ptc therapeutic  
CC (i.e. a small organic molecule that mimics the effect of hedgehog  
CC proteins on patched signalling, or activates or potentiates patched  
CC signalling) in an amount effective to alter the rate of proliferation  
CC of the epithelial cell. The hedgehog therapeutic preferably  
CC comprises at least a bioactive extracellular portion of a hedgehog  
CC protein (see AAY05854-62) encoded by a vertebrate hedgehog gene (see  
CC AAX25617-25), especially a human hedgehog gene. Promotion of  
CC proliferation of epithelial cells can be used to control a wound  
CC healing process in e.g. burn treatment, skin regeneration, skin  
CC grafting, pressure sore treatment, dermal ulcer treatment, post  
CC surgery scar reduction or treatment of ulcerative colitis (claimed).  
CC It can also be used to induce hair growth for the treatment of  
CC alopecia (claimed). Inhibition of the growth of epithelial tissue  
CC can be used to treat or prevent hyperplastic or neoplastic  
CC conditions, e.g. psoriasis, keratosis, acne, comedogenic lesions,  
CC folliculitis and pseudofolliculitis, keratoacanthoma, callositles,  
CC Darier's disease, keloids, hypertrophic scars, or autoimmune  
CC disorders, e.g. aphthous ulcers, pemphigus vulgaris, pemphigus  
CC foliaceus, pemphigus vegetans, pemphigus erythematous,  
CC epidermolysis, lupus lesions, desquamative lesions or carcinomas.  
CC The methods can also be used to counteract the effects of ageing on  
CC skin.  
XX  
XX  
SQ Sequence 425 AA;  
Query Match 100.0%; Score 2218; DB 20; Length 425;  
Best Local Similarity 100.0%; Pred. No. 1.7e-227;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WEMLLRILLYGFCIALVSSGLTGPGRIGRRPKKLTPLAYQFIPNVAEKLIG 60  
DB 1 WEMLLRILLYGFCIALVSSGLTGPGRIGRRPKKLTPLAYQFIPNVAEKLIG 60  
OY 61 ASGREGKITRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOM 120  
DB 61 ASGREGKITRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOM 120  
OY 121 PCVKLRVTEGMEDEGHSEESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYYES 180

```

Db 121 PGVKLRVTEGMDGDHSESLHTEGRAVDITTSDDRSKYGLARLAVEAGFDWYYES 180
Oy 181 KAHHCSVKAENSVAASKGCGPFSATVHLEHGCTKLVDLSPGDRLAADADGRLLYSD 240
Db 181 KAHHCSVKAENSVAASKGCGPFSATVHLEHGCTKLVDLSPGDRLAADADGRLLYSD 240
Oy 241 FLTFIDRMDSSRKLFFYIETROPARLLTLAAHLFVAPDHNSGATSGALFASNV 300
Db 241 FLTFIDRMDSSRKLFFYIETROPARLLTLAAHLFVAPDHNSGATSGALFASNV 300
Oy 301 KEGQRYVYEGEGQQLLPASVHSVSLREBASGAYAPLTAOGTLLINRVLASCAVIEHS 360
Db 301 KEGQRYVYEGEGQQLLPASVHSVSLREBASGAYAPLTAOGTLLINRVLASCAVIEHS 360
Oy 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTTGIMHYSRLYRIGSWVLGDALPLGM 420
Db 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTTGIMHYSRLYRIGSWVLGDALPLGM 420
Oy 421 VAPAS 425
Db 421 VAPAS 425

```

RESULT 3  
AAV05510 standard; Protein; 425 AA.

AAV05510;

05-JUL-1999 (first entry)

Chicken Sonic hedgehog protein Shh.

Sonic hedgehog: Shh protein; chicken: hedgehog therapeutic;  
ptc therapeutic; patched; signal transduction; muscle atrophy;  
cachexia; muscular myopathy; myoblastic sarcoma; therapy.

Gallus sp.

MO9910004-A2.

04-MAR-1999.

28-AUG-1998; 98WO-US17922.

29-AUG-1997; 97US-0057394.

(ONTO-) ONTOGENY INC.

Bladgen CS, Currie PD, Hughes SM, Ingham PW;

WPI: 1999-243557/20.

N-PSDB; AAX25098.

A new method to regulate muscle growth

Disclosure: Page 111-112; 130pp; English.

The present sequence is chicken Sonic hedgehog protein Shh. The  
invention relates to a method for modulating the formation and/or  
maintenance of muscle tissue by ecotopically contacting muscle  
cells, especially muscle stem/progenitor cells, in vitro or in  
vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and  
gene therapy constructs) or ptc therapeutic (i.e. a small organic  
molecule that mimics the effect of hedgehog proteins on patched  
signaling, or activators or potentiators patched signaling) in an  
amount effective to alter the growth state of the treated cells.  
Also claimed is a method for treatment or prevention of disorders  
of, or surgical or cosmetic repair of, such muscle tissues, by  
administering a hedgehog polypeptide or ptc therapeutic. The  
disorder may be muscle atrophy, in particular skeletal muscle  
atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy

CC (all claimed). The hedgehog polypeptide or ptc therapeutic can  
inhibit growth of myoblastic-derived tissue to provide treatment of  
hyperplastic or neoplastic growth of muscle tissue such as in  
myoblastic sarcoma (also claimed). The hedgehog therapeutic  
preferably comprises at least a bioactive extracellular portion of  
a hedgehog protein (see AAV05510-19) encoded by a vertebrate hedgehog  
gene (see AAX25098-107), especially a human hedgehog gene.

Sequence 425 AA;

Query Match 100.0%; Score 2218; DB 20; Length 425;

Best Local Similarity 100.0%; Pred. No. 1.7e-227;

Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 MYEMLLTRLLVGFICALLVSSGLTCGPGRGIGKRRHPKLLPLAYKQIPVAAKRTLG 60
Db 1 MYEMLLTRLLVGFICALLVSSGLTCGPGRGIGKRRHPKLLPLAYKQIPVAAKRTLG 60
Oy 61 ASGRYEGKITRNSERPEKELTPYNNPDIIFKDEENTGADRLMTGRCKDLNALAISVNW 120
Db 61 ASGRYEGKITRNSERPEKELTPYNNPDIIFKDEENTGADRLMTGRCKDLNALAISVNW 120
Oy 121 PGVKLRVTEGMDGDHSESLHTEGRAVDITTSDDRSKYGLARLAVEAGFDWYYES 180
Db 121 PGVKLRVTEGMDGDHSESLHTEGRAVDITTSDDRSKYGLARLAVEAGFDWYYES 180
Oy 181 KAHHCSVKAENSVAASKGCGPFSATVHLEHGCTKLVDLSPGDRLAADADGRLLYSD 240
Db 181 KAHHCSVKAENSVAASKGCGPFSATVHLEHGCTKLVDLSPGDRLAADADGRLLYSD 240
Oy 241 FLTFIDRMDSSRKLFFYIETROPARLLTLAAHLFVAPDHNSGATSGALFASNV 300
Db 241 FLTFIDRMDSSRKLFFYIETROPARLLTLAAHLFVAPDHNSGATSGALFASNV 300
Oy 301 KEGQRYVYEGEGQQLLPASVHSVSLREBASGAYAPLTAOGTLLINRVLASCAVIEHS 360
Db 301 KEGQRYVYEGEGQQLLPASVHSVSLREBASGAYAPLTAOGTLLINRVLASCAVIEHS 360
Oy 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTTGIMHYSRLYRIGSWVLGDALPLGM 420
Db 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTTGIMHYSRLYRIGSWVLGDALPLGM 420
Oy 421 VAPAS 425
Db 421 VAPAS 425

RESULT 4
AAW97765 standard; Protein; 425 AA.
AAW97765;
21-MAY-1999 (first entry)
Chicken Sonic hedgehog (Shh) protein.
Sonic hedgehog: Shh protein; chicken: dopaminergic; GABA-ergic;
ptc therapeutic; patched; signal transduction; Parkinson's disease;
Huntington's disease; amyotrophic lateral sclerosis;
cerebral ischemia; hypoxia; neuroprotective; therapy.
Gallus domesticus.
MO9904775-A2.
04-FEB-1999.
24-JUL-1998; 98WO-US15419.
24-JUL-1997; 97US-0900220.
(ONTO-) ONTOGENY INC.

```





```
|||||
Db 121 PGVKLRVTEGMEDGHSESLHYEGRAVDITTSDRSKYGMALARLAEAGFDWYYES 180
Oy 181 KAHHCYSVAENSVAAKSGCGPGSATVHLEHGTRKLVKDLSPGDRVLAADAGRLYSD 240
Db 181 KAHHCYSVAENSVAAKSGCGPGSATVHLEHGTRKLVKDLSPGDRVLAADAGRLYSD 240
Oy 241 FLTFLDRMDSRKLFYIETROPARALLTTAAHLLEFVAPQHNOSEATGSGQALFASNV 300
Db 241 FLTFLDRMDSRKLFYIETROPARALLTTAAHLLEFVAPQHNOSEATGSGQALFASNV 300
Oy 301 KPGQRYVVLGEGGQQLLPASVHVSLSREASGAVAPLTAGTILINRVLASCYAVIEHS 360
Db 301 KPGQRYVVLGEGGQQLLPASVHVSLSREASGAVAPLTAGTILINRVLASCYAVIEHS 360
Oy 361 WAHMAFAPFRLAQGLLAALCPDGAIPTAATTGTHWYSRLYRIGSWVLDGDLHPFGM 420
Db 361 WAHMAFAPFRLAQGLLAALCPDGAIPTAATTGTHWYSRLYRIGSWVLDGDLHPFGM 420
Oy 421 VAPAS 425
Db 421 VAPAS 425
```

## RESULT 6

```
AA95972
ID AAY95972 standard; Protein; 425 AA.
```

```
XX AC AAY95972;
```

```
DT 05-DEC-2000 (first entry)
```

```
DE Chicken Sonic hedgehog Shh protein.
```

```
XX KW Sonic hedgehog: Shh; chicken; agonist; antagonist;
KW lipid modulator; vacuole; cholesterol disorder; lipid disorder;
KW lipid metabolism; lipid storage; lipid transport; apolipoprotein;
KW triglyceride; hypercholesterolemia; abetalipoproteinemia;
KW hypobetalipoproteinemia; chylomicron retention; Anderson's disease;
KW fat absorption; atherosclerosis; obesity; weight loss;
KW vitamin A disorder; vitamin E disorder; antiilepemia; anorectic;
KW antidiabetic; gene therapy; diagnosis.
```

```
XX OS Gallus sp.
```

```
XX PM MO200051628-A2.
```

```
XX PD 08-SEP-2000.
```

```
XX PF 03-MAR-2000; 2000WO-US05662.
```

```
XX PR 03-MAR-1999; 99US-0122640.
```

```
XX PR 15-MAR-1999; 99US-0124446.
```

```
XX PA (BIOU ) BIOGEN INC.
```

```
XX PI Burkly L, Wang LC;
```

```
XX DR WPI: 2000-611340/58.
```

```
XX DR P-PSDB: AAA50446.
```

```
XX PT Use of lipid modulators (e.g. hedgehog agonists or antagonists) for
XX PT modulating lipid metabolism and storage, especially useful for treating
XX PT lipid metabolism or cholesterol disorders, e.g. obesity or
XX PT hypercholesterolemia -
```

```
XX PS Disclosure: Page 107-109; 136pp; English.
```

```
XX CC The present sequence of that of chicken Sonic hedgehog (Shh)
XX CC protein. The invention provides claimed methods for modulating
XX CC lipid metabolism, for modulating vacuole formation in intestinal
XX CC epithelial cells, for modulating the accumulation of fat in
XX CC intestinal epithelial cells, for treating a cholesterol disorder
```

```
CC and for treating a lipid metabolism disorder in an animal
CC (especially a human) by administering a lipid modulator selected
CC from a hedgehog antagonist or hedgehog agonist. In particular, the
CC lipid metabolism disorder is a lipid storage disorder, a lipid
CC transport disorder, a triglyceride metabolism disorder, a triglyceride
CC disorder, e.g. a triglyceride storage disorder, a triglyceride
CC transport disorder or a triglyceride storage disorder; a
CC diet-induced hypercholesterolemia, hypercholesterolemia,
CC chylomicron-retention disorder, Anderson's disease, a fat
CC absorption disorder, e.g. obesity or associated with weight loss,
CC normotriglyceridemic abetalipoproteinemia, an apolipoprotein-B100
CC deficiency, a fat soluble vitamin disorder, where the fat soluble
CC vitamin is vitamin A or E, or atherosclerosis (all claimed). The
CC hedgehog antagonist binds to the hedgehog receptor, but does not
CC elicit a response. It is preferably a hedgehog mimetic, a modified
CC hedgehog protein, e.g. an inactive hedgehog variant, or a humanised
CC anti-hedgehog homologue, especially a human, chimeric or humanised
CC antibody. The methods are useful in preventing these disorders or
CC protecting a subject from these disorders. The hedgehog antagonist
CC and agonist are also useful in diagnosis and research associated
CC with these disorders. The lipid modulators may also be used as a
CC part of a gene therapy protocol to deliver polynucleotides encoding
CC these lipid modulators.
```

```
XX SQ Sequence 425 AA;
```

```
Query Match 100.0%; Score 2218; DB 21; Length 425;
```

```
Best Local Similarity 100.0%; Pred. No. 1.7e-227; Mismatches 0; Indels 0; Gaps 0;
```

```
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 WVEMLLRLILVGFICALVSSGLTCGPGRGIGRRHPRKLTPLAYKOFINVAEKTIG 60
```

```
Db 1 WVEMLLRLILVGFICALVSSGLTCGPGRGIGRRHPRKLTPLAYKOFINVAEKTIG 60
```

```
Oy 61 ASGRREGKTRNSERFKELTPYNDPIFKDEENTGADRLMTORCKDKINALAISVMNQ 120
```

```
Db 61 ASGRREGKTRNSERFKELTPYNDPIFKDEENTGADRLMTORCKDKINALAISVMNQ 120
```

```
Oy 121 PGVKLRVTEGMEDGHSESLHYEGRAVDITTSDRSKYGMALARLAEAGFDWYYES 180
```

```
Db 121 PGVKLRVTEGMEDGHSESLHYEGRAVDITTSDRSKYGMALARLAEAGFDWYYES 180
```

```
Oy 181 KAHHCYSVAENSVAAKSGCGPGSATVHLEHGTRKLVKDLSPGDRVLAADAGRLYSD 240
```

```
Db 181 KAHHCYSVAENSVAAKSGCGPGSATVHLEHGTRKLVKDLSPGDRVLAADAGRLYSD 240
```

```
Oy 241 FLTFLDRMDSRKLFYIETROPARALLTTAAHLLEFVAPQHNOSEATGSGQALFASNV 300
```

```
Db 241 FLTFLDRMDSRKLFYIETROPARALLTTAAHLLEFVAPQHNOSEATGSGQALFASNV 300
```

```
Oy 301 KPGQRYVVLGEGGQQLLPASVHVSLSREASGAVAPLTAGTILINRVLASCYAVIEHS 360
```

```
Db 301 KPGQRYVVLGEGGQQLLPASVHVSLSREASGAVAPLTAGTILINRVLASCYAVIEHS 360
```

```
Oy 361 WAHMAFAPFRLAQGLLAALCPDGAIPTAATTGTHWYSRLYRIGSWVLDGDLHPFGM 420
```

```
Db 361 WAHMAFAPFRLAQGLLAALCPDGAIPTAATTGTHWYSRLYRIGSWVLDGDLHPFGM 420
```

```
Oy 421 VAPAS 425
```

```
Db 421 VAPAS 425
```

## RESULT 7

```
AA95281
ID AAY95281 standard; Protein; 425 AA.
```

```
XX AC AAY95281;
```

```
XX DT 12-SEP-2000 (first entry)
```

```
XX DE Chicken Sonic hedgehog Shh protein.
```

XX Sonic hedgehog; Shh; chicken; excitotoxicity; Parkinson's disease;  
 KW Huntington's disease; neuronal degeneration; neuroprotective;  
 KM dopaminergic; GABAergic; substantia nigra; therapy.  
 XX  
 OS Gallus sp.  
 XX  
 PN WO200035948-A1.  
 XX  
 PD 22-JUN-2000.  
 XX  
 PF 03-DEC-1999; 99WO-US28721.  
 XX  
 PR 03-DEC-1998; 98WO-US25676.  
 PR 27-JAN-1999; 99US-0238243.  
 PR 03-JUN-1999; 99US-0325602.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 PA (ONTO-) ONTOGENT INC.  
 XX  
 PI Galdes A, Mahanthappa N;  
 XX  
 DR WPI: 2000-431570/37.  
 DR N-PSDB; AAA27876.  
 XX  
 PT Treating disorders involving exotoxicity, e.g. trauma, hypoglycemia,  
 PT senile dementia and Korsakoff's disease, by using lipophilic modified  
 PT hedgehog polypeptide -  
 XX  
 PS Disclosure: Page 149-151; 174pp; English.  
 XX  
 CC The present sequence of that of chicken Sonic hedgehog (Shh)  
 CC protein. The invention relates to a method for promoting the  
 CC survival and/or functional performance of neuronal cells,  
 CC especially substantia nigra, dopaminergic or GABAergic neurons that  
 CC are susceptible to exotoxicity, by contacting the cells, in vitro  
 CC or in vivo, with a lipophilic (e.g. cholesterol) modified hedgehog  
 CC polypeptide. The method is used to treat or prevent Parkinson's  
 CC disease, Huntington's disease, domoic acid poisoning, spinal cord  
 CC trauma, hypoglycemia, mechanical trauma to the nervous system,  
 CC senile dementia, Korsakoff's disease, schizophrenia, AIDS dementia,  
 CC multi-infarct dementia, mood disorders, depression, chemical  
 CC toxicity, neuronal damage associated with uncontrolled seizures  
 CC such as epileptic seizures, neuronal injury associated with HIV and  
 CC AIDS, neurodegeneration associated with Down's syndrome,  
 CC neuropathic pain syndrome, olivopontocerebral atrophy, amyotrophic  
 CC lateral sclerosis, mitochondrial abnormalities, Alzheimer's disease,  
 CC hepatic encephalopathy, Tourette's syndrome and drug addiction (all  
 CC claimed). The lipophilic modified hedgehog polypeptide is also  
 CC useful for promoting survival and/or functional performance of  
 CC neuronal cells susceptible to exotoxicity.  
 XX  
 SQ Sequence 425 AA:  
 Query Match 100.0%; Score 2218; DB 21; Length 425;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-227;  
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 LFLFLDRMSSRKLFYIETROPARLLTPAHLFLFAPQHNQSEAGSTSGALFASNV 300  
 |||||||  
 DB 241 FLFLDRMSSRKLFYIETROPARLLTPAHLFLFAPQHNQSEAGSTSGALFASNV 300  
 QY 301 KPGGRVYLVGEGGQQLPASHVSLSREASGATAPLTAOSTIILINRYLASCYAVIEHS 360  
 |||||||  
 DB 301 KPGGRVYLVGEGGQQLPASHVSLSREASGATAPLTAOSTIILINRYLASCYAVIEHS 360  
 QY 361 WAHNAFAPFRLAOGLLAALCPDGAIPTAATTTGIIHWSRLYRIGSVYLDGDLHPHGM 420  
 |||||||  
 DB 361 WAHNAFAPFRLAOGLLAALCPDGAIPTAATTTGIIHWSRLYRIGSVYLDGDLHPHGM 420  
 QY 421 VAPAS 425  
 |||||  
 DB 421 VAPAS 425  
 RESULT 8  
 ID AAY96243  
 AAAY96243 standard; Protein: 425 AA.  
 XX  
 AC AAY96243;  
 XX  
 DT 11-SEP-2000 (first entry)  
 XX  
 DE Partial chicken Shh.  
 XX  
 KW Chicken; sonic hedgehog; Shh; neuromuscular disorder; neuropathy;  
 KW Guillain-Barre syndrome; peripheral neuropathy; diabetes; alcoholism;  
 KW chronic inflammatory demyelinating polyneuropathy; CIDP;  
 KW gene therapy; infection; inflammation; hereditary neuropathy;  
 KW Charcot-Marie-Tooth disease; vasculitis; lung cancer; tumour;  
 KW multiple myeloma; nutritional imbalance; kidney disease;  
 KW hypothyroid neuropathy; trauma; Refsum's disease; Abetalipoproteinemia;  
 KW Fabry's disease; Krabbe's disease; Metachromatic leukodystrophy;  
 KW Amyotrophic lateral sclerosis; ALS; Miller-Fisher syndrome; amyloidosis;  
 KW hereditary sensory neuropathy Type II; HSN II; B-cell lymphoma;  
 KW Waldenstrom's Macroglobulinemia; Chronic lymphocytic leukaemia;  
 KW neuroprotective; cytoprotective; patched-mediated signal transduction.  
 XX  
 OS Gallus domesticus.  
 XX  
 PN WO200027422-A2.  
 XX  
 PD 18-MAY-2000.  
 XX  
 PF 08-NOV-1999; 99WO-US26334.  
 PF 06-NOV-1998; 98US-0187387.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 PA (ONTO-) ONTOGENT INC.  
 XX  
 PI Galdes A, Mahanthappa N;  
 XX  
 DR WPI: 2000-387341/33.  
 DR N-PSDB; AAA30274.  
 XX  
 PT Novel method of preventing deterioration of peripheral nerves, useful  
 PT for treating or preventing neuropathy, e.g. where associated with  
 PT diabetes or viral infection, by administering hedgehog or patched agent  
 PT -  
 PS Claim 7; Page 127-128; 152pp; English.  
 XX  
 CC The present sequence is the partial chicken sonic hedgehog protein, Shh.  
 CC This sequence inhibits expression of the patched gene which has been  
 CC implicated in neuromuscular disorders (neuropathies). This sequence may  
 CC therefore be used for treating neuromuscular disorders i.e. preventing  
 CC degradation in function of motor or sensory nerves and protecting  
 CC peripheral nerve cells under conditions that normally cause neuropathy.

CC A variety of neuromuscular disorders may be treated: Gullain-Barre  
 CC syndrome, GBS; peripheral neuropathy; diabetic neuropathy;  
 CC alcohol-induced neuropathy; chronic inflammatory demyelinating  
 CC polyneuropathy, CIPD; infection-induced neuropathy, including HIV  
 CC infection; inflammation-induced neuropathy; hereditary neuropathy e.g.  
 CC Charcot-Marie-Tooth disease (CMT), Familial Amyloidotic neuropathy,  
 CC Reissum's disease, Abetalipoproteinemia, Tanager disease, Krabbe's  
 CC disease, Metachromatic leukodystrophy, Fabry's disease, Dejerine-Sottas  
 CC syndrome, Hereditary sensory neuropathy Type II (HSN II) and Amyotrophic  
 CC lateral sclerosis (ALS); acute neuropathy e.g. Miller-Fisher syndrome;  
 CC neuropathy caused by vasculitis; neuropathy associated with tumours e.g.  
 CC lung cancer, multiple myeloma, B-cell lymphoma, Waldenstrom's  
 CC Macroglobulinaemia, Chronic Lymphocytic Leukaemia; neuropathy associated  
 CC with: amyloidosis, nutritional imbalance, kidney disease, trauma; and  
 CC hypothyroid neuropathy. The coding sequence may be used in gene therapy  
 CC of the above disorders.  
 CC  
 XX Sequence 425 AA:  
 SQ  
 Query Match 100.0%; Score 2218; DB 21; Length 425;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-227;  
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVEMLLTTRILLVGFICALLVSSGLTCGPGRGIGKRRHPRKLTPLAYKQFIPNVAEKTIG 60  
 DB 1 MVEMLLTTRILLVGFICALLVSSGLTCGPGRGIGKRRHPRKLTPLAYKQFIPNVAEKTIG 60  
 QY ASGRYEGKITRNSERKELTPNNYNDITFKDEENTGADRLMTORCKDKLAALISVNNOW 120  
 DB ASGRYEGKITRNSERKELTPNNYNDITFKDEENTGADRLMTORCKDKLAALISVNNOW 120  
 QY 121 PGVKLRVTEGMDGDHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYES 180  
 DB 121 PGVKLRVTEGMDGDHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYES 180  
 QY 121 PGVKLRVTEGMDGDHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYES 180  
 DB 121 PGVKLRVTEGMDGDHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYES 180  
 QY 181 KAHHCSVKAENSVAAKSGCGPFGSATVHLEHGCTLVKDLSPGDRVLAADADGRLLYSD 240  
 DB 181 KAHHCSVKAENSVAAKSGCGPFGSATVHLEHGCTLVKDLSPGDRVLAADADGRLLYSD 240  
 QY 181 KAHHCSVKAENSVAAKSGCGPFGSATVHLEHGCTLVKDLSPGDRVLAADADGRLLYSD 240  
 DB 181 KAHHCSVKAENSVAAKSGCGPFGSATVHLEHGCTLVKDLSPGDRVLAADADGRLLYSD 240  
 QY 241 FLTFIDRMDSSRKLTFYIETROPARALLTAAHLTFVAPOHNSEATGSGALPASNV 300  
 DB 241 FLTFIDRMDSSRKLTFYIETROPARALLTAAHLTFVAPOHNSEATGSGALPASNV 300  
 QY 301 KEGQRYVVLGEGGQQLLPASVHVSLEESAAGYAPLTAQGTILLINRVLASCYAVIEHS 360  
 DB 301 KEGQRYVVLGEGGQQLLPASVHVSLEESAAGYAPLTAQGTILLINRVLASCYAVIEHS 360  
 QY 361 WAHMAFAPRLAOGLLAALCPDGAIPTAATTGTHMYSRLYRIGSWLDGALHPLGM 420  
 DB 361 WAHMAFAPRLAOGLLAALCPDGAIPTAATTGTHMYSRLYRIGSWLDGALHPLGM 420  
 QY 421 VAPAS 425  
 DB 421 VAPAS 425

RESULT 9  
 AAY70676  
 ID AAY70676 standard; Protein; 425 AA.  
 XX  
 AC AAY70676;  
 XX  
 DT 18-JUL-2000 (first entry)  
 XX  
 DE Chicken Sonic hedgehog (Shh) protein.  
 XX  
 KW Sonic hedgehog; Shh; chicken; growth modulator; therapeutic agent;  
 KW lung; hedgehog; hh; patched; ptc; fibroblast growth factor; fgf-10;  
 KW anti-proliferative; anticancer; vulnereary; antineumatic; hypotensive;  
 KW anti-inflammatory; antisthmatic; antiarthritic; tuberculostatic; asthma;  
 KW antimicrobial; antiallergy; treatment; prevention; lung diseases; cancer;  
 KW cystic fibrosis; bronchitis; emphysema; respiratory distress syndrome;  
 KW tuberculosis; wound healing; lung transplantation.

XX  
 OS Gallus domesticus.  
 XX  
 PN WO200015246-A2.  
 XX  
 PD 23-MAR-2000.  
 XX  
 PF 10-SEP-1999; 99WO-US20500.  
 XX  
 PR 11-SEP-1998; 98US-0099952.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Pepliceall C, Lewis P, McMahon AP;  
 XX  
 DR WPI: 2000-271252/23.  
 XX  
 N-PSDB; AA52257.  
 XX  
 PT Modulation of lung tissue or cell growth rate used for treating or  
 PT preventing damage to lung tissue comprises ectopically contacting  
 PT tissue with hedgehog therapeutic, patched therapeutic or fibroblast  
 PT growth factor-10  
 XX  
 PS Claim 14; Page 112-114; 143pp; English.  
 XX  
 CC The patent discloses a method for modulating the growth state of  
 CC epithelial or mesenchymal cells of the lung, by ectopically altering the  
 CC the tissue with a therapeutic agent, that can effectively alter the rate  
 CC of proliferation of cells. This agent can be selected from hedgehog (hh),  
 CC patched (ptc) or fibroblast growth factor (fgf)-10 therapeutics. It  
 CC involves a direct or indirect antagonism of patched-mediated regulation  
 CC of gene expression. This method is useful for the treatment or prevention  
 CC of lung diseases, like cancer, cystic fibrosis, bronchopneumocystosis,  
 CC bronchitis, bronchospasm, sarcoidosis, silicosis, eosinophilic granuloma,  
 CC ankylosing spondylitis, emphysema, tuberculosis, respiratory distress  
 CC syndrome, allergic rhinitis, asthma, pulmonary fibrosis and primary  
 CC pulmonary hypertension. It is also used to control wound healing or other  
 CC reformation processes in the lung and augment lung transplantation. The  
 CC present sequence is the chicken sonic hedgehog (Shh) protein, essential  
 CC for development of the respiratory system. Hedgehog polypeptides can be  
 CC used to control the formation and/or maintenance of the lung tissue.  
 CC  
 XX Sequence 425 AA:  
 SQ  
 Query Match 100.0%; Score 2218; DB 21; Length 425;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-227;  
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVEMLLTTRILLVGFICALLVSSGLTCGPGRGIGKRRHPRKLTPLAYKQFIPNVAEKTIG 60  
 DB 1 MVEMLLTTRILLVGFICALLVSSGLTCGPGRGIGKRRHPRKLTPLAYKQFIPNVAEKTIG 60  
 QY 61 ASGRYEGKITRNSERKELTPNNYNDITFKDEENTGADRLMTORCKDKLAALISVNNOW 120  
 DB 61 ASGRYEGKITRNSERKELTPNNYNDITFKDEENTGADRLMTORCKDKLAALISVNNOW 120  
 QY 121 PGVKLRVTEGMDGDHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYES 180  
 DB 121 PGVKLRVTEGMDGDHSESLHYEGRAVDITTSDDRSKYGMARLAVEAGFDWYYES 180  
 QY 181 KAHHCSVKAENSVAAKSGCGPFGSATVHLEHGCTLVKDLSPGDRVLAADADGRLLYSD 240  
 DB 181 KAHHCSVKAENSVAAKSGCGPFGSATVHLEHGCTLVKDLSPGDRVLAADADGRLLYSD 240  
 QY 241 FLTFIDRMDSSRKLTFYIETROPARALLTAAHLTFVAPOHNSEATGSGALPASNV 300  
 DB 241 FLTFIDRMDSSRKLTFYIETROPARALLTAAHLTFVAPOHNSEATGSGALPASNV 300  
 QY 301 KEGQRYVVLGEGGQQLLPASVHVSLEESAAGYAPLTAQGTILLINRVLASCYAVIEHS 360  
 DB 301 KEGQRYVVLGEGGQQLLPASVHVSLEESAAGYAPLTAQGTILLINRVLASCYAVIEHS 360  
 QY 361 WAHMAFAPRLAOGLLAALCPDGAIPTAATTGTHMYSRLYRIGSWLDGALHPLGM 420

|||||  
Db 361 WAHMAFAFRLAAGLLALCPDGAIPATAITTTGTHWTSRLLYRIGSVNLDDALHPLGM 420  
QY 421 VAPAS 425  
|||||  
Db 421 VAPAS 425  
RESULT 10  
AAG65743  
ID AAG65743 standard; Protein: 425 AA.  
XX  
AC AAG65743;  
XX  
DT 07-JAN-2002 (first entry)  
XX  
DE Chicken sonic hedgehog (Shh) polypeptide.  
XX  
KW Adipocyte; hedgehog polypeptide; desert hedgehog; indian hedgehog; Dh; Ih; sonic hedgehog; Shh; therapeutic; cytostatic; tumor.  
XX  
OS Gallus sp.  
XX  
PN MO200164238-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-US06450.  
XX  
PR 29-FEB-2000; 2000US-186058P.  
XX  
PA (CURT-) CURIS INC.  
XX  
PI Zehentner B, Leser-Reiff U, Bartscher H;  
XX  
DR WPI: 2001-607352/69.  
DR N-PSDB; AA166771.  
XX  
PT Method for regulating formation and/or maintenance of adipocyte tissue  
PT by contacting pre-adipocyte or adipocyte cells with a hedgehog  
PT polypeptide or plc therapeutic -  
XX  
PS Disclosure; Page 90-92; 132pp; English.  
XX  
CC The invention provides a method for regulating formation and/or  
CC maintenance of adipocyte tissue that comprises contacting pre adipocyte  
CC or adipocyte cells with a hedgehog polypeptide or plc therapeutic. The  
CC method is used for regulating the growth state of an adipocyte stem/  
CC progenitor cell, and treating or preventing disorders of, or surgical or  
CC cosmetic repair of, adipocyte tissues, e.g. for treating or preventing  
CC hyperplastic or neoplastic conditions affecting adipocyte tissue, such  
CC as soft tissue tumors, especially adipose cell tumors, e.g. lipomas,  
CC fibrolipomas, lipoblastomas, lipomatosis, fibromas, hemangiomas and/or  
CC liposarcomas. Hedgehog polypeptides can be used in combination with other  
CC therapeutic agents. The present sequence represents a chicken sonic  
CC hedgehog (Shh) polypeptide.  
XX  
SQ Sequence 425 AA;  
Query Match 100.0%; Score 2218; DB 22; Length 425;  
Best Local Similarity 100.0%; Pred. No. 1,7e-227;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVEMLLRLILVGFICALVSSGLTCGPGNGICRRHPPKLTPLAVYQFIPTNVAEKTIG 60  
|||||  
Db 1 MVEMLLRLILVGFICALVSSGLTCGPGNGICRRHPPKLTPLAVYQFIPTNVAEKTIG 60  
QY 61 ASGREGKITSNREKRLTNYNDITFKDEENTGADRLMTQKDKLNLALSYVMQW 120  
|||||  
Db 61 ASGREGKITSNREKRLTNYNDITFKDEENTGADRLMTQKDKLNLALSYVMQW 120  
QY 121 PGVKLRVEGMDGDGHHSEESLHYEGRVADITTSRDRSKTGMLARLAVEAGFDVYTES 180  
|||||

Db 121 PGVKLRVEGMDGDGHHSEESLHYEGRVADITTSRDRSKTGMLARLAVEAGFDVYTES 180  
QY 181 KAHICSVKAENSVAKSGGCEPGSATVHLEHGTRKLVKDSLPGDRLVAADAGRLYSD 240  
|||||  
Db 181 KAHICSVKAENSVAKSGGCEPGSATVHLEHGTRKLVKDSLPGDRLVAADAGRLYSD 240  
QY 241 FLTFIDRNDSSKRLFYETROPARLLITAAHLFVAPQHNQSEATSTSQALFASNV 300  
|||||  
Db 241 FLTFIDRNDSSKRLFYETROPARLLITAAHLFVAPQHNQSEATSTSQALFASNV 300  
QY 301 KPGQRYVYLGEQGQOLLASVHVSLSREASGAVAPLTAOGTILNRYLASCYAVTEHS 360  
|||||  
Db 301 KPGQRYVYLGEQGQOLLASVHVSLSREASGAVAPLTAOGTILNRYLASCYAVTEHS 360  
QY 361 WAHMAFAFRLAAGLLALCPDGAIPATAITTTGTHWTSRLLYRIGSVNLDDALHPLGM 420  
|||||  
Db 361 WAHMAFAFRLAAGLLALCPDGAIPATAITTTGTHWTSRLLYRIGSVNLDDALHPLGM 420  
QY 421 VAPAS 425  
|||||  
Db 421 VAPAS 425  
RESULT 11  
AAB85733  
ID AAB85733 standard; Protein: 425 AA.  
XX  
AC AAB85733;  
XX  
DT 29-OCT-2001 (first entry)  
XX  
DE Chicken sonic hedgehog (Shh) polypeptide.  
XX  
KW Hedgehog protein; sonic hedgehog; Shh; indian hedgehog; Ih; Dh;  
KW desert hedgehog; cell differentiation; chicken.  
XX  
OS Gallus sp.  
XX  
PN US6271363-B1.  
XX  
PD 07-AUG-2001.  
XX  
PF 20-OCT-1997; 97US-0954698.  
XX  
PR 05-JUN-1995; 95US-0462386.  
PR 30-DEC-1993; 93US-0176427.  
PR 14-DEC-1994; 94US-0356060.  
PR 04-MAY-1995; 95US-0435093.  
XX  
PA (HARD ) HARVARD COLLEGE.  
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
XX  
PI Ingham PW, McMahon AP, Tablin CJ;  
XX  
DR WPI: 2001-456723/49.  
DR N-PSDB; AAH76107.  
XX  
PT Novel nucleic acid encoding a hedgehog polypeptide, used to produce the  
PT polypeptide, which is used to promote proliferation, survival, and/or  
PT differentiation of neuronal and mesodermal tissue -  
XX  
PS Claim 1; Column 117-120; 118pp; English.  
XX  
CC The invention relates to nucleic acids encoding hedgehog proteins  
CC selected from sonic hedgehog (Shh), indian hedgehog (Ihh), desert  
CC hedgehog (Dhh) polypeptides. The hedgehog genes are involved in the  
CC formation of ordered spatial arrangements of differentiated tissue in  
CC vertebrates. The nucleic acid sequences are useful for producing hedgehog  
CC proteins, used for promoting differentiation of, or survival of  
CC differentiated, neuronal cells, and for promoting proliferation, survival  
CC or differentiation of mesenchymal, endodermal or ectodermal tissue,  
CC particularly chondrocytes, or testicular germ line cells. The present  
CC sequence represents a chicken Shh polypeptide.

[illegible]

XX	PT	Modulating immune function comprises administration of a hedgehog or	XX
PT	ptc agent, for treating e.g. diabetes, eczematous dermatitis, urticaria		
PT	or vasculitis -		
XX			
PS	Claim 4; Page 77-78; 105bp; English.		
CC	The present sequence represents a hedgehog (Shh) polypeptide. Hedgehog		
CC	gene products and signal transduction pathways involving hedgehog are		
CC	involved in the maturation of T lymphocytes. The specification describes		
CC	a method for modulating immune function, by administration of a hedgehog		
CC	or patched (ptc) polypeptide, agonists or antagonists. The method is		
CC	used to treat disorders affecting the regulation of lymphocytes,		
CC	particularly maturation and/or activation of T lymphocytes. It is used		
CC	to treat bacterial or viral infection, diabetes, nutritional		
CC	deficiencies, graft rejection or other hyperacute response such as		
CC	kidney, heart, lung, bone marrow spleen skin or cornea transplant or		
CC	autoimmune disorders such as multiple sclerosis, psoriasis or atopic		
CC	dermatitis. The method is used to treat inflammatory, proliferative and		
CC	hyperproliferative diseases, as well as cutaneous manifestations of		
CC	immunological disorders such as eczematous dermatitis, urticaria,		
CC	vasculitis and scleroderma.		
XX			
XX	Sequence 425 AA;		
Query Match	100.0%; Score 2218; DB 22; Length 425;		
Best Local Similarity	100.0%; Pred. No. 1,7e-227;		
Matches	425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 MVEMLLRLILVGFICALVSSGLTSGPGRGIGRRHPKLTPLAYKQFIPNVAEKTG 60		
Db	1 MVEMLLRLILVGFICALVSSGLTSGPGRGIGRRHPKLTPLAYKQFIPNVAEKTG 60		
QY	61 ASGRGEGKITRNSERFKELTPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNW 120		
Db	61 ASGRGEGKITRNSERFKELTPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNW 120		
QY	121 PGVKLRVTEGDEDEHSHSESLHYEGRAVDITTSDDRSKYGMLARLVAEAGFDWYYES 180		
Db	121 PGVKLRVTEGDEDEHSHSESLHYEGRAVDITTSDDRSKYGMLARLVAEAGFDWYYES 180		
QY	181 KAHHTCSYKANSVNAKSGGCEPGSATVNLHEGGRKLYKDSLPGDRVLAADADGRLLYSD 240		
Db	181 KAHHTCSYKANSVNAKSGGCEPGSATVNLHEGGRKLYKDSLPGDRVLAADADGRLLYSD 240		
QY	241 FLTFLDRLDSSRKLFYVETROPARALLTFAHLLFVAPOHNSSEATSTSGQALFASN 300		
Db	241 FLTFLDRLDSSRKLFYVETROPARALLTFAHLLFVAPOHNSSEATSTSGQALFASN 300		
QY	301 KPGQRYVYVGBGGQOLLPAHSVSVSLREBASGAVAPLTAOGTILINRYLASCYAVIEEHS 360		
Db	301 KPGQRYVYVGBGGQOLLPAHSVSVSLREBASGAVAPLTAOGTILINRYLASCYAVIEEHS 360		
QY	361 WAHNAFAFPRLAOGLLALCPDGAIPPTATTTGTGHHVSRLLYRIGSNWLDGDALHPGM 420		
Db	361 WAHNAFAFPRLAOGLLALCPDGAIPPTATTTGTGHHVSRLLYRIGSNWLDGDALHPGM 420		
QY	421 VAPAS 425		
Db	421 VAPAS 425		
RESULT 13			
AAED04682			
ID	AAED04682 standard; Protein; 425 AA.		
XX	AAED04682;		
XX	AAED04682;		
XX	04-SEP-2001 (first entry)		
XX	Chicken sonic hedgehog (Shh) protein.		
XX	Chicken: hedgehog protein; nontropic; neuroprotective; anticonvulsant;		

KW cytostatic; therapy; Alzheimer's disease; Parkinson's disease; injury;  
KW Huntington's chorea; amyotrophic lateral sclerosis; multiple sclerosis  
KW nervous system aging; neurodegenerative disease; immunological disease  
KW malignant glioma; medulloblastoma; neuroectodermal tumour; cancer;  
KW extracellular signalling protein.

OS Gallus sp.

PN WO200134654-A1.

PD 17-MAY-2001.

PF 02-NOV-2000; 2000WO-US30405.

PR 05-NOV-1999; 99US-0164025.

PA (BIOJ ) BIOGEN INC.

PI Strauch K;

DR WPI; 2001-329075/34.

XX

PT Novel isolated hedgehog fusion polypeptide useful for treating  
PT neurological conditions such as Alzheimer's disease, Parkinson's  
PT disease, Huntington's chorea, amyotrophic lateral sclerosis, and  
PT multiple sclerosis -

PS Disclosure; Page 111-112; 178bp; English.

The present invention relates to hedgehog fusion proteins. Hedgehog proteins are a family of extracellular signalling proteins that regulate various aspects of embryonic development both in vertebrates and in invertebrates. Hedgehog fusion protein is useful for the prophylaxis or treatment of any condition or disease state for which a hedgehog or patched protein constituent is efficacious and in the diagnosis of CC constituents or conditions of disease states in biological system or specimens and for diagnostic purposes in non-physiological systems. Hedgehog fusion protein is useful for treating neurological conditions due to injury, aging of nervous system, including Alzheimer's disease, chronic neurodegenerative diseases of the nervous system, including Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis and chronic immunological diseases of nervous system including multiple sclerosis and malignant gliomas, medulloblastomas, neuroectodermal tumours and to specifically target medical therapies against cancers and tumours which express the receptor for the protein. The present sequence is chicken sonic hedgehog (Shh) protein.

**SQ Sequence 425 AA;**

Query Match	100.0%	Score 2218;	DB 22;	Length 425;
Best Local Similarly	100.0%;	Pred. No. 1.7e-227;		
Matches 425; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MVEMLILTRILLYVETICALLVSSGLTGGPGGIGKRRHPPKILPRLAYAKORLIPNVAEITLG	60
Db	1	MVEMLILTRILLYVETICALLVSSGLTGGPGGIGKRRHPPKILPRLAYAKORLIPNVAEITLG	60
QY	61	ASGREYEGKITNSENSEFKELPAPNPNPDIIFDKDEENTGADRLMTORCKOKLMAIAISVNMOM	120
Db	61	ASGREYEGKITNSENSEFKELPAPNPNPDIIFDKDEENTGADRLMTORCKOKLMAIAISVNMOM	120
QY	121	PGVRLPRTTEGMDDEGHHSSESLHIEGNAVDITTSDDRRSKTYGMLARLAVENAGDFMWYYES	180
Db	121	PGVRLPRTTEGMDDEGHHSSESLHIEGNAVDITTSDDRRSKTYGMLARLAVENAGDFMWYYES	180
QY	181	KAHLHCISVKAENSVAAKSGGCFPCSAIVHLEHGSTKLVKDIJSPCDRYLADAOGRILLYSD	240
Db	181	KAHLHCISVKAENSVAAKSGGCFPCSAIVHLEHGSTKLVKDIJSPCDRYLADAOGRILLYSD	240
QY	241	FLVFLDRMDSRKLFVYIETROPARLLTPAHLFLVAPOHNSSEATGTSQALFEASNV	3000
Db	241	FLVFLDRMDSRKLFVYIETROPARLLTPAHLFLVAPOHNSSEATGTSQALFEASNV	3000

QY	301	KPGQNVYVLGEGGQQLPASHVSHVSLREASGATAPLTAOCTILINNVLASCAVIEHS	360
Db	301	KPGQNVYVLGEGGQQLPASHVSHVSLREASGATAPLTAOCTILINNVLASCAVIEHS	360
QY	361	WAHNAFAPFRLAOGILALCPDGAIPPTAATTTTGIHWYSLRLYRIGSWVLDDGALHPLGM	420
Db	361	WAHNAFAPFRLAOGILALCPDGAIPPTAATTTTGIHWYSLRLYRIGSWVLDDGALHPLGM	420
QY	421	VAPAS 425	
Db	421	VAPAS 425	

RESULT 14  
22F06377

AAE05372 standard; Protein; 425 AA

... AC AAE05372

DT 12-SEP-2001 (first entry)

DE Chicken Sonic hedgehog protein.

KM Chicken Sonic hedgehog; Shh; morphogenetic signal; neuron;  
KM embryonic patterning; cell culture; cell differentiation; ischaemia;  
KM cell proliferative disorder; Intracerebral grafting; Huntington's chorea  
KM neurological disorder; Alzheimer's disease; Parkinson's disease;  
KM amyotrophic lateral sclerosis; ALS; multiple sclerosis.

OS Gallus sp.

Key	Location/Qualifiers
FH	1..26
FT	/label= Signal-peptide
FT	27..425
FT	/label= Mature-Shh-protein

PN US6261786-B1

PD 17-JUL-2001.

PF 02-JUL-1996; 96US-0674509.

PR 30-DEC-1993; 93US-0176427.

PR 04-MAY-1995; 95US-0435093.

PR 05-JUN-1995; 95US-0462386.

PA (IMCR) IMPERIAL CANCER RES

XX

XX

DR N-PSDB; AAD10146.

PT Screening compounds

PT contacting polypeptide

XX

2000

the present invention relates to assay for screening compounds that potentiate or inhibit binding of hedgehog polypeptide to naturally occurring patched receptor. The hedgehog proteins comprise morphogenic signals produced by embryonic patterning centres, and are involved in the formation and maintenance of ordered spatial arrangements of differentiated tissues in vertebrates, both adult and embryonic. The proteins can be used to generate and/or maintain an array of different vertebrate tissues both in vitro and in vivo. The invention also relates

CC to a method for modulating growth, differentiation or survival of a  
 CC mammalian cell (e.g. neuron, testicular cell) responsive to hedgehog  
 CC induction. Hedgehog agonists and antagonists can be used in cell culture  
 CC techniques to enhance survival and maintenance of neurons and various  
 CC vertebrate organogenic pathways. The hedgehog gene is useful in  
 CC determining whether a patient is at the risk of disorder characterised by  
 CC unwanted cell proliferation or aberrant control of differentiation. The  
 CC hedgehog proteins or mimetics can be used to induce foetal neurons  
 CC especially neuronal stem cells in intracerebral grafting. The protein  
 CC or its mimetic can be used in the treatment of neurological conditions  
 CC e.g. injury to nervous system, ischaemia resulting from stroke,  
 CC Alzheimer's disease, Parkinson's disease, Huntington's chorea,  
 CC amyotrophic lateral sclerosis (ALS) and multiple sclerosis. The present  
 CC sequence is chicken Sonic hedgehog (Shh) protein.

SQ Sequence 425 AA:

Query Match 100.0%; Score 2218; DB 22; Length 425;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-227;

Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MVEMLLTRILLVGFICALLVSSGLTCGPGRGIGKRHPKKLPLAYKOFIPNVAEKTIG 60  
 DB 1 MVEMLLTRILLVGFICALLVSSGLTCGPGRGIGKRHPKKLPLAYKOFIPNVAEKTIG 60  
 OY 61 ASGRYEGKITRNSERFKEKLPNNYNDIIFKDEENTGADRLMTORCKKLNALAISVNNQW 120  
 DB 61 ASGRYEGKITRNSERFKEKLPNNYNDIIFKDEENTGADRLMTORCKKLNALAISVNNQW 120  
 OY 121 PGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWYYES 180  
 DB 121 PGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWYYES 180  
 OY 181 KAHHCYSKAKENSVAAKSGCGPESATVHLEHGRTKLVKRLSPGDRVLADADGRLLYSD 240  
 DB 181 KAHHCYSKAKENSVAAKSGCGPESATVHLEHGRTKLVKRLSPGDRVLADADGRLLYSD 240  
 OY 241 FLTFLDMDSSRKLFFYIETROPARALLTAHLLFVAPQHNSEATGSGALFASNV 300  
 DB 241 FLTFLDMDSSRKLFFYIETROPARALLTAHLLFVAPQHNSEATGSGALFASNV 300  
 OY 301 KPGQRYVVLGEGGOQLLPASVHSVLSREASGAYAPLTAOCTILINRVLASCAVIEHS 360  
 DB 301 KPGQRYVVLGEGGOQLLPASVHSVLSREASGAYAPLTAOCTILINRVLASCAVIEHS 360  
 OY 361 MAHMAFAPFRLAAGLLAALCPDAIPTAATTGTHWYSRLYRIGSWVLDGDLHPFLGM 420  
 DB 361 MAHMAFAPFRLAAGLLAALCPDAIPTAATTGTHWYSRLYRIGSWVLDGDLHPFLGM 420  
 OY 421 VAPAS 425  
 DB 421 VAPAS 425

RESULT 15

AAB31217 AAB31217 standard; Protein: 425 AA.

AC AAB31217;

DT 20-APR-2001 (first entry)

XX Amino acid sequence of chicken sonic hedgehog protein (Shh).

XX Hedgehog related-protein; sonic hedgehog protein; Shh; Ischemia; stroke;

XX desert hedgehog protein; Dhh; Indian hedgehog protein; Ihh; neuron;

XX neurological condition; nervous system injury; tumour-induced injury;

XX aging; Alzheimer's disease; chronic neurodegenerative disease;

XX Parkinson's disease; Huntington's chorea; amyotrophic lateral sclerosis;

XX spinocerebellar degeneration; chronic immunological disease;

XX multiple sclerosis.

XX US6165747-A.  
 PN 26-DEC-2000.  
 XX  
 XX 05-JUN-1995; 95US-0460900.  
 PR 30-DEC-1993; 93US-0176427.  
 PR 14-DEC-1994; 94US-0356060.  
 PR 04-MAY-1995; 95US-0435093.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
 XX  
 XX Ingham FM, McMahon AP, Tablin CJ, Marti-gorostiza E, Bumcrot DA;  
 DR N-PSDB; AAC87074.  
 XX  
 PT Polynucleotides encoding hedgehog proteins, useful for treating  
 PT diseases of nervous system such as Alzheimer's disease, Parkinson's  
 PT disease, Huntington's chorea, amyotrophic lateral sclerosis, multiple  
 PT sclerosis -

Claim 10; Columns 135-138; 119pp; English.

The present sequence represents a hedgehog related-protein. The  
 CC specification describes a sonic hedgehog protein (Shh), a desert  
 CC hedgehog protein (Dhh), and an Indian hedgehog protein (Ihh). The  
 CC hedgehog polynucleotides are useful in diagnostic, in antisense  
 CC therapy and in therapeutic assays for detecting and treating disorders  
 CC involving, e.g., aberrant expression of vertebrate hedgehog homologue.  
 CC Hedgehog polypeptides are useful therapeutically to enhance survival  
 CC of neurons and other neuron cells and in treating neurological  
 CC conditions deriving from acute, subacute, or chronic injury to the  
 CC nervous system, including traumatic injury, chemical injury, vessel  
 CC injury and deficits (such as the ischemia resulting from stroke),  
 CC together with infectious/inflammatory and induced-induced injury, aging  
 CC of the nervous system including Alzheimer's disease, chronic  
 CC neurodegenerative diseases of the nervous system, including Parkinson's  
 CC disease, Huntington's chorea, amyotrophic lateral sclerosis,  
 CC spinocerebellar degenerations, and chronic immunological diseases of  
 CC the nervous system or affecting the nervous system, including multiple  
 CC sclerosis.

SQ Sequence 425 AA:

Query Match 100.0%; Score 2218; DB 22; Length 425;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-227;

Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MVEMLLTRILLVGFICALLVSSGLTCGPGRGIGKRHPKKLPLAYKOFIPNVAEKTIG 60  
 DB 1 MVEMLLTRILLVGFICALLVSSGLTCGPGRGIGKRHPKKLPLAYKOFIPNVAEKTIG 60  
 OY 61 ASGRYEGKITRNSERFKEKLPNNYNDIIFKDEENTGADRLMTORCKKLNALAISVNNQW 120  
 DB 61 ASGRYEGKITRNSERFKEKLPNNYNDIIFKDEENTGADRLMTORCKKLNALAISVNNQW 120  
 OY 121 PGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWYYES 180  
 DB 121 PGVKLRVTEGWDGHHSESLHYEGRAVDITTSDDRSKYGMRLARLAVAGFDWYYES 180  
 OY 181 KAHHCYSKAKENSVAAKSGCGPESATVHLEHGRTKLVKRLSPGDRVLADADGRLLYSD 240  
 DB 181 KAHHCYSKAKENSVAAKSGCGPESATVHLEHGRTKLVKRLSPGDRVLADADGRLLYSD 240  
 OY 241 FLTFLDMDSSRKLFFYIETROPARALLTAHLLFVAPQHNSEATGSGALFASNV 300  
 DB 241 FLTFLDMDSSRKLFFYIETROPARALLTAHLLFVAPQHNSEATGSGALFASNV 300  
 OY 301 KPGQRYVVLGEGGOQLLPASVHSVLSREASGAYAPLTAOCTILINRVLASCAVIEHS 360  
 DB 301 KPGQRYVVLGEGGOQLLPASVHSVLSREASGAYAPLTAOCTILINRVLASCAVIEHS 360



Db 301 KPGORVYVLGEGGQULPASVHSLSREASGAVAPLTAOGTILINRVLASCYAVIEHS 360  
Oy 361 WAHNAFAPFRLAQGLLALCPDGAIPATAATTTGTHWYSRLIYRIGSWVLDGDALHPLGM 420  
Db 361 WAHNAFAPFRLAQGLLALCPDGAIPATAATTTGTHWYSRLIYRIGSWVLDGDALHPLGM 420  
Oy 421 VAPAS 425  
Db 421 VAPAS 425

Search completed: February 20, 2003, 10:10:17  
Job time : 32.834 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:08:38 ; Search time 11.1257 Seconds  
(without alignments)  
1123.956 Million cell updates/sec

Title: US-09-827-110A-10

Perfect score: 2218  
Sequence: 1 MVEMLLTRILVGFICALL.....GSWVLDGDLHPLGNVAPAS 425

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A.COMB.pdp:\*  
2: /cgn2\_6/prodata/1/1aa/5B.COMB.pdp:\*  
3: /cgn2\_6/prodata/1/1aa/6A.COMB.pdp:\*  
4: /cgn2\_6/prodata/1/1aa/6B.COMB.pdp:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS.COMB.pdp:\*  
6: /cgn2\_6/prodata/1/1aa/Dackfilest1.pdp:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2218	100.0	425	1	US-08-176-427B-2
2	2218	100.0	425	2	US-08-356-060A-8
3	2218	100.0	425	4	US-08-460-900C-8
4	2218	100.0	425	4	US-08-674-509B-8
5	2218	100.0	425	4	US-08-954-698-8
6	2218	100.0	425	4	US-08-957-874-8
7	2218	100.0	425	4	US-09-325-256-17
8	2085	94.0	425	3	US-08-946-329A-19
9	2085	94.0	425	4	US-08-567-357A-19
10	2085	94.0	425	4	US-08-729-743A-19
11	2085	94.0	425	4	US-08-349-498-19
12	2085	94.0	425	5	PCT-US95-15463-19
13	2085	94.0	425	5	PCT-US95-15923-19
14	1807	81.5	462	1	US-08-748-591-4
15	1807	81.5	462	1	US-08-748-591-9
16	1807	81.5	475	2	US-08-356-060A-13
17	1807	81.5	475	4	US-08-460-900C-13
18	1807	81.5	475	4	US-08-674-509B-13
19	1807	81.5	475	4	US-08-954-698-13
20	1807	81.5	475	4	US-08-957-874-13
21	1807	81.5	475	4	US-09-325-256-22
22	1791	80.7	437	3	US-08-946-329A-20
23	1791	80.7	437	4	US-08-567-357A-20
24	1791	80.7	437	4	US-08-729-743A-20
25	1791	80.7	437	4	US-09-057-860A-6
26	1791	80.7	437	4	US-08-349-498-20
27	1791	80.7	437	4	US-09-293-505-14

28	1791	80.7	437	5	PCT-US95-15463-20	Sequence 20, Appl
29	1791	80.7	437	5	PCT-US95-15923-20	Sequence 20, Appl
30	1788	80.6	437	1	US-08-176-427B-8	Sequence 8, Appl
31	1788	80.6	437	2	US-08-356-060A-11	Sequence 11, Appl
32	1788	80.6	437	4	US-08-460-900C-11	Sequence 11, Appl
33	1788	80.6	437	4	US-08-674-509B-11	Sequence 11, Appl
34	1788	80.6	437	4	US-08-954-698-11	Sequence 11, Appl
35	1788	80.6	437	4	US-08-957-874-11	Sequence 11, Appl
36	1788	80.6	437	4	US-09-325-256-20	Sequence 20, Appl
37	1780	80.3	437	4	US-08-757-230A-2	Sequence 2, Appl
38	1780	80.3	437	4	US-08-757-230A-9	Sequence 9, Appl
39	1780	80.3	437	5	PCT-US95-02315-2	Sequence 2, Appl
40	1494	67.4	418	3	US-08-946-329A-18	Sequence 18, Appl
41	1494	67.4	418	4	US-08-567-357A-18	Sequence 18, Appl
42	1494	67.4	418	4	US-08-729-743A-18	Sequence 18, Appl
43	1494	67.4	418	4	US-08-757-230A-7	Sequence 7, Appl
44	1494	67.4	418	4	US-08-349-498-18	Sequence 18, Appl
45	1494	67.4	418	5	PCT-US95-15463-18	Sequence 18, Appl

#### ALIGNMENTS

```
RESULT 1
US-08-176-427B-2
; Sequence 2, Application US/08176427B
; Patent No. 5789543
;
GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; NUMBER OF INVENTIONS: Proteins and Uses Related Thereto
;
CORRESPONDENCE ADDRESSES: 33
; ADDRESSSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/176,427B
; FILING DATE: 30-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
;
INFORMATION FOR SEQ. ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-176-427B-2
;
Query Match 100.0%; Score 2218; DB 1; Length 425;
Best Local Similarity 100.0%; Pred. No. 1.5e-241;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 1 MVEMLLTRILVGFICALLVSSGLTCGPGRCIGKRRHPRKLLPLVAYKQIPVNAEKTIG 60
Db 1 MVEMLLTRILVGFICALLVSSGLTCGPGRCIGKRRHPRKLLPLVAYKQIPVNAEKTIG 60
```

```

Oy 61 ASGREGKTRNSERFELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNQ 120
    |||||||
Db 61 ASGREGKTRNSERFELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNQ 120
Oy 121 PGVKRLVTEGMEDEGHSEESLHYEGRAVDITTSRDRSKYGMRLAVEGFDMVYYES 180
    |||||||
Db 121 PGVKRLVTEGMEDEGHSEESLHYEGRAVDITTSRDRSKYGMRLAVEGFDMVYYES 180
Oy 181 KAHHCYKAENSVAAKSGGCGPGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLYSD 240
    |||||||
Db 181 KAHHCYKAENSVAAKSGGCGPGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLYSD 240
Oy 241 FLTFIDRMDSSKRLFYVETROPARLLITAAHLFVAPQHNQSEATGSGQALFASNV 300
    |||||||
Db 241 FLTFIDRMDSSKRLFYVETROPARLLITAAHLFVAPQHNQSEATGSGQALFASNV 300
Oy 301 KPGQRYVYLGEGGQQLLPASVHVSRLREASGAVAPLTAOGTILINRYLASCYAVIEHS 360
    |||||||
Db 301 KPGQRYVYLGEGGQQLLPASVHVSRLREASGAVAPLTAOGTILINRYLASCYAVIEHS 360
Oy 361 WAHMAFAFRLAOGLLALCPDGAIPTAATTTTGGIHWYSRLLYRIGSVWLDGDLHPJGM 420
    |||||||
Db 361 WAHMAFAFRLAOGLLALCPDGAIPTAATTTTGGIHWYSRLLYRIGSVWLDGDLHPJGM 420
Oy 421 VAPAS 425
    |||||
Db 421 VAPAS 425

```

## RESULT 2

US-08-356-060A-8  
Sequence (8) Application US/08356060A  
Patent No. 5844079

## GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356, 060A  
FILING DATE: 14-DEC-1994  
CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMT-006CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-356-060A-8

Query Match 100.0%; Score 2218; DB 2; Length 425;  
Best Local Similarity 100.0%; Pred. No. 1,5e-241;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 MEMLLTRILLVGFICALYSSGLTCGPGRGIGRRRPPKILPLAYAQFIPNVAEKTIG 60
    |||||||
Db 1 MEMLLTRILLVGFICALYSSGLTCGPGRGIGRRRPPKILPLAYAQFIPNVAEKTIG 60
Oy 61 ASGREGKTRNSERFELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNQ 120
    |||||||
Db 61 ASGREGKTRNSERFELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNQ 120
Oy 121 PGVKRLVTEGMEDEGHSEESLHYEGRAVDITTSRDRSKYGMRLAVEGFDMVYYES 180
    |||||||
Db 121 PGVKRLVTEGMEDEGHSEESLHYEGRAVDITTSRDRSKYGMRLAVEGFDMVYYES 180
Oy 181 KAHHCYKAENSVAAKSGGCGPGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLYSD 240
    |||||||
Db 181 KAHHCYKAENSVAAKSGGCGPGSATVHLEHGTRKLVKDLSPGDRVLAADADGRLYSD 240
Oy 241 FLTFIDRMDSSKRLFYVETROPARLLITAAHLFVAPQHNQSEATGSGQALFASNV 300
    |||||||
Db 241 FLTFIDRMDSSKRLFYVETROPARLLITAAHLFVAPQHNQSEATGSGQALFASNV 300
Oy 301 KPGQRYVYLGEGGQQLLPASVHVSRLREASGAVAPLTAOGTILINRYLASCYAVIEHS 360
    |||||||
Db 301 KPGQRYVYLGEGGQQLLPASVHVSRLREASGAVAPLTAOGTILINRYLASCYAVIEHS 360
Oy 361 WAHMAFAFRLAOGLLALCPDGAIPTAATTTTGGIHWYSRLLYRIGSVWLDGDLHPJGM 420
    |||||||
Db 361 WAHMAFAFRLAOGLLALCPDGAIPTAATTTTGGIHWYSRLLYRIGSVWLDGDLHPJGM 420
Oy 421 VAPAS 425
    |||||
Db 421 VAPAS 425

```

## RESULT 3

US-08-460-900C-8  
Sequence 8, Application US/08460900C  
Patent No. 6165747

## GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
APPLICANT: Bumcrot, David A.  
APPLICANT: Marti-Gorostiza, Elisa  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HONG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460, 900C  
FILING DATE: 5-JUNE-1995  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/435,093  
FILING DATE: 4-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356, 060  
FILING DATE: 14-DEC-1994  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-460-900C-8

Query Match 100.0%; Score 2218; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 1.5e-241;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVEMLLTLLILVGFICALLVSSGLTCGPGRGIGKRRHPKLLPLAYKOFIPNVAEKTIG 60  
DB 1 MVEMLLTLLILVGFICALLVSSGLTCGPGRGIGKRRHPKLLPLAYKOFIPNVAEKTIG 60  
QY 61 ASGRYEGKITRNSERFELTPNPNPDIIFKDEENTGADRLMTORCKDLNALAISVNW 120  
DB 61 ASGRYEGKITRNSERFELTPNPNPDIIFKDEENTGADRLMTORCKDLNALAISVNW 120  
QY 121 PGKLVATEGMDGDHSESLHYEGRAVDITTSDDRSTYGMALRLAVEGPDWYYES 180  
DB 121 PGKLVATEGMDGDHSESLHYEGRAVDITTSDDRSTYGMALRLAVEGPDWYYES 180  
QY 181 KAHHCYSKAEVNSVAASGCGFPGSATVHLEHGSTKLVKDLSPGDRVLAADADGRLLYSD 240  
DB 181 KAHHCYSKAEVNSVAASGCGFPGSATVHLEHGSTKLVKDLSPGDRVLAADADGRLLYSD 240  
QY 241 FLTFLDMDSSSKLFYIETROPARLLLTAAHLLEFYAPOHNOSEATGSGALFASNV 300  
DB 241 FLTFLDMDSSSKLFYIETROPARLLLTAAHLLEFYAPOHNOSEATGSGALFASNV 300  
QY 301 KPGORVYVLEGGQQLLPASVHSVLSREASGAYAPLTAGCTILINVLASCAVIEHS 360  
DB 301 KPGORVYVLEGGQQLLPASVHSVLSREASGAYAPLTAGCTILINVLASCAVIEHS 360  
QY 361 MAHMAFAPFLAAGLLAALCPDGAIPTAATTGIIHWSRLYRIGSWLDGDLHPHGM 420  
DB 361 MAHMAFAPFLAAGLLAALCPDGAIPTAATTGIIHWSRLYRIGSWLDGDLHPHGM 420  
QY 421 VAPAS 425  
DB 421 VAPAS 425

RESULT 4  
US-08-674-509B-8  
Sequence 8, Application US/08674509B  
Patent No. 6261786  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
APPLICANT: Marijo, Valeria  
TITLE OF INVENTION: SCREENING ASSAYS FOR HEDGEHOG AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/674,509B  
FILING DATE: 02-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/460,900  
FILING DATE: 05-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.06  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-674-509B-8

Query Match 100.0%; Score 2218; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 1.5e-241;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVEMLLTLLILVGFICALLVSSGLTCGPGRGIGKRRHPKLLPLAYKOFIPNVAEKTIG 60  
DB 1 MVEMLLTLLILVGFICALLVSSGLTCGPGRGIGKRRHPKLLPLAYKOFIPNVAEKTIG 60  
QY 61 ASGRYEGKITRNSERFELTPNPNPDIIFKDEENTGADRLMTORCKDLNALAISVNW 120  
DB 61 ASGRYEGKITRNSERFELTPNPNPDIIFKDEENTGADRLMTORCKDLNALAISVNW 120  
QY 121 PGKLVATEGMDGDHSESLHYEGRAVDITTSDDRSTYGMALRLAVEGPDWYYES 180  
DB 121 PGKLVATEGMDGDHSESLHYEGRAVDITTSDDRSTYGMALRLAVEGPDWYYES 180  
QY 181 KAHHCYSKAEVNSVAASGCGFPGSATVHLEHGSTKLVKDLSPGDRVLAADADGRLLYSD 240  
DB 181 KAHHCYSKAEVNSVAASGCGFPGSATVHLEHGSTKLVKDLSPGDRVLAADADGRLLYSD 240  
QY 241 FLTFLDMDSSSKLFYIETROPARLLLTAAHLLEFYAPOHNOSEATGSGALFASNV 300  
DB 241 FLTFLDMDSSSKLFYIETROPARLLLTAAHLLEFYAPOHNOSEATGSGALFASNV 300  
QY 301 KPGORVYVLEGGQQLLPASVHSVLSREASGAYAPLTAGCTILINVLASCAVIEHS 360  
DB 301 KPGORVYVLEGGQQLLPASVHSVLSREASGAYAPLTAGCTILINVLASCAVIEHS 360  
QY 361 MAHMAFAPFLAAGLLAALCPDGAIPTAATTGIIHWSRLYRIGSWLDGDLHPHGM 420  
DB 361 MAHMAFAPFLAAGLLAALCPDGAIPTAATTGIIHWSRLYRIGSWLDGDLHPHGM 420  
QY 421 VAPAS 425  
DB 421 VAPAS 425

RESULT 5  
US-08-954-698-8  
Sequence 8, Application US/08954698  
Patent No. 6271363  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tabin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

TITLE OF INVENTION: Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HONG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/954,698  
FILING DATE: 20-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/462,386  
FILING DATE: 05-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 04-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-954-698-8

Query Match 100.0%; Score 2218; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 1.5e-241;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WEMLLLRILLVGFTICALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60  
DB 1 WEMLLLRILLVGFTICALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60  
QY 61 ASGREGKITRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVMNQ 120  
DB 61 ASGREGKITRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVMNQ 120  
QY 121 PGVKLRVTEGMDEGHSEESLHYEGRAVDITTSDRSKSYGMRLARLAVEAGFDWVYYS 180  
DB 121 PGVKLRVTEGMDEGHSEESLHYEGRAVDITTSDRSKSYGMRLARLAVEAGFDWVYYS 180  
QY 161 KAHHCSTKAENSVAAKSGGCGPGSATYVLEHGRTKLYKVDLSPGRVLAADADGRLISD 240  
DB 161 KAHHCSTKAENSVAAKSGGCGPGSATYVLEHGRTKLYKVDLSPGRVLAADADGRLISD 240  
QY 241 FLTFIDRSDSKLFYVETROPARLLTLTAHLTFVAPOHQSEATSGTSGOALFASNV 300  
DB 241 FLTFIDRSDSKLFYVETROPARLLTLTAHLTFVAPOHQSEATSGTSGOALFASNV 300  
QY 301 KPGQRYVYVLEGEGQQLLPASVHSVLSREASGAYAPLTAQGITLLINRYLASCYAVIEHS 360  
DB 301 KPGQRYVYVLEGEGQQLLPASVHSVLSREASGAYAPLTAQGITLLINRYLASCYAVIEHS 360  
QY 361 WAHMAFAPRLAOGILALCPDGAIP7AATTTTG1HWYSRLLYRIGSVWLDGDLHP1GM 420

DB 361 WAHMAFAPRLAOGILALCPDGAIP7AATTTTG1HWYSRLLYRIGSVWLDGDLHP1GM 420  
QY 421 VAPAS 425  
DB 421 VAPAS 425

RESULT 6  
US-08-957-874-8  
Sequence 8, Application US/08957874  
Patent No. 6384192  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tablin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
TITLE OF INVENTION: Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HONG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/957,874  
FILING DATE: 20-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/462,386  
FILING DATE: 5-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 4-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.09  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-957-874-8

Query Match 100.0%; Score 2218; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 1.5e-241;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WEMLLLRILLVGFTICALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60  
DB 1 WEMLLLRILLVGFTICALVSSGLTCGPGRGIGRRRHPKLTPLAYKQFIPNVAEKTIG 60  
QY 61 ASGREGKITRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVMNQ 120  
DB 61 ASGREGKITRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVMNQ 120

QY 121 PGYKLRVTEGMDDEGHHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVAGFDWYYES 180  
DB 121 PGYKLRVTEGMDDEGHHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVAGFDWYYES 180  
QY 181 KAHHCVCVKAENSVAAASGGCFPGSATVHLEHGSTKLYVKLSPDRVLADADGRLLYSD 240  
DB 181 KAHHCVCVKAENSVAAASGGCFPGSATVHLEHGSTKLYVKLSPDRVLADADGRLLYSD 240  
QY 241 FLFLFDMDSSRKLFFYIETROPARLLLTAAHLLEVAPOHNOSEATGTSQGLFASNV 300  
DB 241 FLFLFDMDSSRKLFFYIETROPARLLLTAAHLLEVAPOHNOSEATGTSQGLFASNV 300  
QY 301 KPGORVYVLEGGQQLLPASVHSVSLREASGAYAPLTAOCTILINRVLASCYAVIEHS 360  
DB 301 KPGORVYVLEGGQQLLPASVHSVSLREASGAYAPLTAOCTILINRVLASCYAVIEHS 360  
QY 361 MAHMAFAPFLAOGLLAALCPDGAIPPAATTTGIIHWSRLYRIGSWVLDGDLHPLGM 420  
DB 361 MAHMAFAPFLAOGLLAALCPDGAIPPAATTTGIIHWSRLYRIGSWVLDGDLHPLGM 420  
QY 421 VAPAS 425  
DB 421 VAPAS 425

RESULT 7  
US-09-325-256-17  
; Sequence 17, Application US/09325256  
; Patent No. 6444793  
; GENERAL INFORMATION:  
; APPLICANT: PEPINSKY, R. BLAKE  
; APPLICANT: BAKER, DARREN P.  
; APPLICANT: WEN, DINGYI  
; APPLICANT: WILLIAMS, KEVIN P.  
; APPLICANT: GARGER, ELLEN A.  
; APPLICANT: TAYLOR, FREDERICK R.  
; APPLICANT: GALDES, ALPHONSE  
; APPLICANT: PORTER, JEFFREY  
; TITLE OF INVENTION: HYDROPHOBICALLY-MODIFIED PROTEIN COMPOSITIONS AND  
; TITLE OF INVENTION: METHODS  
; FILE REFERENCE: BIV-067.01  
; CURRENT APPLICATION NUMBER: US/09/325,256  
; CURRENT FILING DATE: 1999-06-03  
; PRIOR APPLICATION NUMBER: 60/099,800  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/078,935  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/089,685  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/067,423  
; PRIOR FILING DATE: 1997-12-03  
; PRIOR APPLICATION NUMBER: PCT/US98/25676  
; PRIOR FILING DATE: 1998-12-03  
; NUMBER OF SEQ. ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ. ID NO: 17  
; LENGTH: 425  
; TYPE: PRT  
; ORGANISM: Gallus sp.  
US-09-325-256-17

Query Match 100.0%; Score 2218; DB 4; Length 425;  
Best Local Similarity 100.0%; Pred. No. 1.5e-241;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVEKLLTRILLVGFALLVSSSLTGPRGRGIGKRHRPKKLTPLAKOKTIPNVAEKTIG 60  
DB 1 MVEKLLTRILLVGFALLVSSSLTGPRGRGIGKRHRPKKLTPLAKOKTIPNVAEKTIG 60  
QY 61 ASGRYEGKTRNSERFKELTPNYPDITTFKDEENTGADRLMTORCKKLAALATISVNW 120  
DB 61 ASGRYEGKTRNSERFKELTPNYPDITTFKDEENTGADRLMTORCKKLAALATISVNW 120

QY 121 PGYKLRVTEGMDDEGHHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVAGFDWYYES 180  
DB 121 PGYKLRVTEGMDDEGHHSESLHYEGRAVDITTSDRDRSKYGMRLARLAVAGFDWYYES 180  
QY 181 KAHHCVCVKAENSVAAASGGCFPGSATVHLEHGSTKLYVKLSPDRVLADADGRLLYSD 240  
DB 181 KAHHCVCVKAENSVAAASGGCFPGSATVHLEHGSTKLYVKLSPDRVLADADGRLLYSD 240  
QY 241 FLFLFDMDSSRKLFFYIETROPARLLLTAAHLLEVAPOHNOSEATGTSQGLFASNV 300  
DB 241 FLFLFDMDSSRKLFFYIETROPARLLLTAAHLLEVAPOHNOSEATGTSQGLFASNV 300  
QY 301 KPGORVYVLEGGQQLLPASVHSVSLREASGAYAPLTAOCTILINRVLASCYAVIEHS 360  
DB 301 KPGORVYVLEGGQQLLPASVHSVSLREASGAYAPLTAOCTILINRVLASCYAVIEHS 360  
QY 361 MAHMAFAPFLAOGLLAALCPDGAIPPAATTTGIIHWSRLYRIGSWVLDGDLHPLGM 420  
DB 361 MAHMAFAPFLAOGLLAALCPDGAIPPAATTTGIIHWSRLYRIGSWVLDGDLHPLGM 420  
QY 421 VAPAS 425  
DB 421 VAPAS 425

RESULT 8  
US-08-946-329A-19  
; Sequence 19, Application US/08946329A  
; Patent No. 6057091  
; GENERAL INFORMATION:  
; APPLICANT: Beachy, Philip A.  
; APPLICANT: Porter, Jeffrey A.  
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
; NUMBER OF SEQUENCES: 109  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/946,329A  
; FILING DATE: 07-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/061,323  
; FILING DATE: 07-OCT-1996  
; APPLICATION NUMBER: 08/7729,743  
; FILING DATE: 10-JUL-1996  
; APPLICATION NUMBER: 08/567,357  
; FILING DATE: 04-DEC-1995  
; APPLICATION NUMBER: 08/349,498  
; FILING DATE: 02-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halle, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/140001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; INFORMATION FOR SEQ. ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 425 amino acids  
; TYPE: amino acid  
; TOPOLOGY: both  
; MOLECULE TYPE: protein  
US-08-946-329A-19

Query Match 94.0%; Score 2085; DB 3; Length 425;  
Best Local Similarity 96.2%; Pred. No. 1.6e-226;  
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

1 MEWMLLRILLVGFICALVSSGLTCGPGRGIGRRHHPKKTLPAYKQFIPNVAEKTIG 60  
1 MEWMLLRILLVGFICALVSSGLTCGPGRGIGRRHHPKKTLPAYKQFIPNVAEKTIG 60  
61 ASGREGKTRNSERFKELIPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNQ 120  
61 ASGREGKTRNSERFKELIPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNQ 120  
121 PGVKLRVTEGMDEGDHSESLHYEGRAVDITTSRDRSKYGMRLARLAVAGFDVYYES 180  
121 PGVKLRVTEGMDEGDHSESLHYEGRAVDITTSRDRSKYGMRLARLAVAGFDVYYES 180  
181 KAHICSVKAENSVAKSGGCGPFSATYVHLEHGTRKLVKDSLHGDRVLAADADGRLLVSD 240  
181 KAHICSVKAENSVAKSGGCGPFSATYVHLEHGTRKLVKDSLHGDRVLAADADGRLLVSD 240  
241 F-LTFELDRMDSRRKLFYVETROPARLLTAAHLFFVAPQHNSEATGTSQGLFASN 299  
241 F-LTFELDRMDSRRKLFYVETROPARLLTAAHLFFVAPQHNSEATGTSQGLFASN 299  
300 VPGQRYVYLGEQGQQLPASVHSVSLREASGAYAPLTAOGTILINRYLASCYAVIEEH 359  
300 VPGQRYVYLGEQGQQLPASVHSVSLREASGAYAPLTAOGTILINRYLASCYAVIEEH 359  
360 SWAHAFAPFRLAOGLLAALCPDGAIPTAATTTGIIHWSRLYRIGSVWLDGDLHPLG 419  
360 SWAHAFAPFRLAOGLLAALCPDGAIPTAATTTGIIHWSRLYRIGSVWLDGDLHPLG 419  
420 MWAPAS 425  
420 MWAPAS 425

RESULT 9  
US-08-567-357A-19  
Sequence 19, Application US/08567357A  
Patent No. 6132728  
GENERAL INFORMATION:  
APPLICANT: Beachy, Philip A.  
APPLICANT: Moon, Randall T.  
APPLICANT: Porter, Jeffrey A.  
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/567,357A  
FILING DATE: 04-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/349,498  
FILING DATE: 02-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/080001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: both  
MOLECULE TYPE: protein  
US-08-567-357A-19

Query Match 94.0%; Score 2085; DB 4; Length 425;  
Best Local Similarity 96.2%; Pred. No. 1.6e-226;  
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

1 MEWMLLRILLVGFICALVSSGLTCGPGRGIGRRHHPKKTLPAYKQFIPNVAEKTIG 60  
1 MEWMLLRILLVGFICALVSSGLTCGPGRGIGRRHHPKKTLPAYKQFIPNVAEKTIG 60  
61 ASGREGKTRNSERFKELIPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNQ 120  
61 ASGREGKTRNSERFKELIPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNQ 120  
121 PGVKLRVTEGMDEGDHSESLHYEGRAVDITTSRDRSKYGMRLARLAVAGFDVYYES 180  
121 PGVKLRVTEGMDEGDHSESLHYEGRAVDITTSRDRSKYGMRLARLAVAGFDVYYES 180  
181 KAHICSVKAENSVAKSGGCGPFSATYVHLEHGTRKLVKDSLHGDRVLAADADGRLLVSD 240  
181 KAHICSVKAENSVAKSGGCGPFSATYVHLEHGTRKLVKDSLHGDRVLAADADGRLLVSD 240  
241 F-LTFELDRMDSRRKLFYVETROPARLLTAAHLFFVAPQHNSEATGTSQGLFASN 299  
241 F-LTFELDRMDSRRKLFYVETROPARLLTAAHLFFVAPQHNSEATGTSQGLFASN 299  
300 VPGQRYVYLGEQGQQLPASVHSVSLREASGAYAPLTAOGTILINRYLASCYAVIEEH 359  
300 VPGQRYVYLGEQGQQLPASVHSVSLREASGAYAPLTAOGTILINRYLASCYAVIEEH 359  
360 SWAHAFAPFRLAOGLLAALCPDGAIPTAATTTGIIHWSRLYRIGSVWLDGDLHPLG 419  
360 SWAHAFAPFRLAOGLLAALCPDGAIPTAATTTGIIHWSRLYRIGSVWLDGDLHPLG 419  
420 MWAPAS 425  
420 MWAPAS 425

RESULT 10  
US-08-729-743A-19  
Sequence 19, Application US/08729743A  
Patent No. 6214794  
GENERAL INFORMATION:  
APPLICANT: Beachy, Philip A.  
APPLICANT: Moon, Randall T.  
APPLICANT: Porter, Jeffrey A.  
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/729,743A  
FILING DATE: 07-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/567,357  
FILING DATE: 04-DEC-1995  
APPLICATION NUMBER: 08/349,498



FILING DATE: 02-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hallie, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/099001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: both  
MOLECULE TYPE: protein  
US-08-729-743A-19

Query Match 94.0%; Score 2085; DB 4; Length 425;  
Best Local Similarity 96.2%; Pred. No. 1.6e-226;  
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

QY 1 MVELLTLRLVGFICALVSSGLTGPCRGIGKRRHPKRLPLAYKQFIPNVAEKTIG 60  
DB 1 MVELLTLRLVGFICALVSSGLTGPCRGIGKRRHPKRLPLAYKQFIPNVAEKTIG 60  
QY 61 ASGRYEGKITRNSEREKELTPNTNPDIIFKDEENTGADRLMTORCKDKLNALISVNCW 120  
DB 61 ASGRYEGKITRNSEREKELTPNTNPDIIFKDEENTGADRLMTORCKDKLNALISVNCW 120  
QY 121 PGVLRTEGMDGHHSESLHYEGRAVDITSDRDSKYGMARLAVAGCDWYYES 180  
DB 121 PGVLRTEGMDGHHSESLHYEGRAVDITSDRDSKYGMARLAVAGCDWYYES 180  
QY 181 KAHIHCSKAENSVAAKSGCFFGSATVHLEHGKTKLVKDLSPGDRVLADADGRLLYSD 240  
DB 181 KAHIHCSKAENSVAAKSGCFFGSATVHLEHGKTKLVKDLSPGDRVLADADGRLLYSD 240  
QY 241 F-LTFDRLDMSRKLFYVETROPARLLTTAAHLFVAPQHNOSEATGTSQALFASN 299  
DB 241 F-LTFDRLDMSRKLFYVETROPARLLTTAAHLFVAPQHNOSEATGTSQALFASN 299  
QY 300 VKPGQRYVVLGEGGQQLLPASVSVSLREBSAGVAPLTAGTLLNRVLAASCVATEEH 359  
DB 300 VKPGQRYVVLGEGGQQLLPASVSVSLREBSAGVAPLTAGTLLNRVLAASCVATEEH 359  
QY 360 SMAHMAFAPRLAOGILAAICPDGAIPTAATTTGIIHWSRLLYRIGSWVLDDGALHPLG 419  
DB 360 SMAHMAFAPRLAOGILAAICPDGAIPTAATTTGIIHWSRLLYRIGSWVLDDGALHPLG 419  
QY 420 MVAPAS 425  
DB 420 MVAPAS 425

RESULT 11  
US-08-349-498-19  
Sequence 19, Application US/08349498  
Patent No. 6281332  
GENERAL INFORMATION:  
APPLICANT: Beachy, Philip A.  
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/349,498  
FILING DATE: 02-DEC-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Hallie, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/043001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: both  
MOLECULE TYPE: protein  
US-08-349-498-19

Query Match 94.0%; Score 2085; DB 4; Length 425;  
Best Local Similarity 96.2%; Pred. No. 1.6e-226;  
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

QY 1 MVELLTLRLVGFICALVSSGLTGPCRGIGKRRHPKRLPLAYKQFIPNVAEKTIG 60  
DB 1 MVELLTLRLVGFICALVSSGLTGPCRGIGKRRHPKRLPLAYKQFIPNVAEKTIG 60  
QY 61 ASGRYEGKITRNSEREKELTPNTNPDIIFKDEENTGADRLMTORCKDKLNALISVNCW 120  
DB 61 ASGRYEGKITRNSEREKELTPNTNPDIIFKDEENTGADRLMTORCKDKLNALISVNCW 120  
QY 121 PGVLRTEGMDGHHSESLHYEGRAVDITSDRDSKYGMARLAVAGCDWYYES 180  
DB 121 PGVLRTEGMDGHHSESLHYEGRAVDITSDRDSKYGMARLAVAGCDWYYES 180  
QY 181 KAHIHCSKAENSVAAKSGCFFGSATVHLEHGKTKLVKDLSPGDRVLADADGRLLYSD 240  
DB 181 KAHIHCSKAENSVAAKSGCFFGSATVHLEHGKTKLVKDLSPGDRVLADADGRLLYSD 240  
QY 241 F-LTFDRLDMSRKLFYVETROPARLLTTAAHLFVAPQHNOSEATGTSQALFASN 299  
DB 241 F-LTFDRLDMSRKLFYVETROPARLLTTAAHLFVAPQHNOSEATGTSQALFASN 299  
QY 300 VKPGQRYVVLGEGGQQLLPASVSVSLREBSAGVAPLTAGTLLNRVLAASCVATEEH 359  
DB 300 VKPGQRYVVLGEGGQQLLPASVSVSLREBSAGVAPLTAGTLLNRVLAASCVATEEH 359  
QY 360 SMAHMAFAPRLAOGILAAICPDGAIPTAATTTGIIHWSRLLYRIGSWVLDDGALHPLG 419  
DB 360 SMAHMAFAPRLAOGILAAICPDGAIPTAATTTGIIHWSRLLYRIGSWVLDDGALHPLG 419  
QY 420 MVAPAS 425  
DB 420 MVAPAS 425

RESULT 12  
PCT-US95-15463-19  
Sequence 19, Application PC/TUS9515463  
GENERAL INFORMATION:  
APPLICANT: The Johns Hopkins University School of Medicine  
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/15463  
FILING DATE: 01-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.  
REGISTRATION NUMBER: 38,347  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: both  
MOLECULE TYPE: protein  
PCT-US95-15463-19

Query Match 94.0%; Score 2085; DB 5; Length 425;  
Best Local Similarity 96.2%; Pred. No. 1.6e-226;  
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

QY 1 MVEMLLTPLLVGFCALVSSGLTCGPGRGIGRRHHPKLTPLAYKOFINVAEKTG 60  
DB 1 MVEMLLTPLLVGFCALVSSGLTCGPGRGIGRRHHPKLTPLAYKOFINVAEKTG 60  
QY 61 ASGREGKTRNSERKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQ 120  
DB 61 ASGREGKTRNSERKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQ 120  
QY 121 PGVKLRVTGMDGDHSHSESLHYEGRAVDITTSDBRSKYGMRLAVEAGFDVYTES 180  
DB 121 PGVKLRVTGMDGDHSHSESLHYEGRAVDITTSDBRSKYGMRLAVEAGFDVYTES 180  
QY 181 KAHICSVKAENSVAAKSGCGPFSATVHLEHGTRKLVKDSLPGDRVLAADAGRLVSD 240  
DB 181 KAHICSVKAENSVAAKSGCGPFSATVHLEHGTRKLVKDSLPGDRVLAADAGRLVSD 240  
QY 241 F-LTFLDRDSSRKLFYIETROPARALLTAHLLFVAPQHNOSEATGSGALFASN 299  
DB 241 F-LTFLDRDSSRKLFYIETROPARALLTAHLLFVAPQHNOSEATGSGALFASN 299  
QY 300 VPRGQRYVVLGGGQQLLPASVHSVSLREASGATAPLTAOCTILINRYLASCYAVIEEH 359  
DB 300 VPRGQRYVVLGGGQQLLPASVHSVSLREASGATAPLTAOCTILINRYLASCYAVIEEH 359  
QY 360 SWAHNAFAPFRLAOGILALCPDGAIPPAATTTTGIIHWSRLLYRIGSVLDGDLHPHG 419  
DB 360 SWAHNAFAPFRLAOGILALCPDGAIPPAATTTTGIIHWSRLLYRIGSVLDGDLHPHG 419  
QY 420 MYAPAS 425  
DB 420 MYAPAS 425

## RESULT 13

PCT-US95-15923-19  
Sequence 19, Application PC/TUS9515923  
GENERAL INFORMATION:  
APPLICANT: The Johns Hopkins University School of Medicine, et al.  
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA

COUNTRY: U.S.A.  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/15923  
FILING DATE: 04-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/043W01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: both  
MOLECULE TYPE: protein  
PCT-US95-15923-19

Query Match 94.0%; Score 2085; DB 5; Length 425;  
Best Local Similarity 96.2%; Pred. No. 1.6e-226;  
Matches 410; Conservative 1; Mismatches 13; Indels 2; Gaps 2;

QY 1 MVEMLLTPLLVGFCALVSSGLTCGPGRGIGRRHHPKLTPLAYKOFINVAEKTG 60  
DB 1 MVEMLLTPLLVGFCALVSSGLTCGPGRGIGRRHHPKLTPLAYKOFINVAEKTG 60  
QY 61 ASGREGKTRNSERKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQ 120  
DB 61 ASGREGKTRNSERKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQ 120  
QY 121 PGVKLRVTGMDGDHSHSESLHYEGRAVDITTSDBRSKYGMRLAVEAGFDVYTES 180  
DB 121 PGVKLRVTGMDGDHSHSESLHYEGRAVDITTSDBRSKYGMRLAVEAGFDVYTES 180  
QY 181 KAHICSVKAENSVAAKSGCGPFSATVHLEHGTRKLVKDSLPGDRVLAADAGRLVSD 240  
DB 181 KAHICSVKAENSVAAKSGCGPFSATVHLEHGTRKLVKDSLPGDRVLAADAGRLVSD 240  
QY 241 F-LTFLDRDSSRKLFYIETROPARALLTAHLLFVAPQHNOSEATGSGALFASN 299  
DB 241 F-LTFLDRDSSRKLFYIETROPARALLTAHLLFVAPQHNOSEATGSGALFASN 299  
QY 300 VPRGQRYVVLGGGQQLLPASVHSVSLREASGATAPLTAOCTILINRYLASCYAVIEEH 359  
DB 300 VPRGQRYVVLGGGQQLLPASVHSVSLREASGATAPLTAOCTILINRYLASCYAVIEEH 359  
QY 360 SWAHNAFAPFRLAOGILALCPDGAIPPAATTTTGIIHWSRLLYRIGSVLDGDLHPHG 419  
DB 360 SWAHNAFAPFRLAOGILALCPDGAIPPAATTTTGIIHWSRLLYRIGSVLDGDLHPHG 419  
QY 420 MYAPAS 425  
DB 420 MYAPAS 425

## RESULT 14

US-08-748-591-4  
Sequence 4, Application US/08748591  
Patent No. 5759811  
GENERAL INFORMATION:  
APPLICANT: Epstein, Ervin  
APPLICANT: Hu, Zhilan  
APPLICANT: Bonifas, Jeanette  
TITLE OF INVENTION: Mutant Human Hedgehog Gene

NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish and Richardson  
STREET: 2200 Sand Hill Road  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/748,591  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J  
REGISTRATION NUMBER: 36,677  
REFERENCE/DOCKET NUMBER: 06510/067001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 322-5070  
TELEFAX: (415) 854-0875  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-748-591-4

Query Match 81.5%; Score 1807; DB 1; Length 462;  
Best Local Similarity 77.8%; Pred. No. 4,7e-195;  
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

DB 4 MLLTRLLVGFICALVSSGLTCGPGRGIGKRRHPKRLPLAYKOPIPVAEKTIGASG 63  
1 MLLARCLLVLVSSLYCGLACGPGRGFGKRRHPKRLPLAYKOPIPVAEKTIGASG 60  
QY 64 RYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SYNWMPGV 123  
DB 61 RYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SYNWMPGV 120  
QY 124 KLRTGEMDDGHHSESLHYEGRAVDITTSDDRKYGMILARLAVAGFDWYYESKAH 183  
DB 121 KLRTGEMDDGHHSESLHYEGRAVDITTSDDRKYGMILARLAVAGFDWYYESKAH 180  
QY 184 IHC5VKAENSVAAKSGCGPFGSATVHLEHGTRKLVKDLSPGDRVLADADGRLLYSDFLT 243  
DB 181 IHC5VKAENSVAAKSGCGPFGSATVHLEHGTRKLVKDLSPGDRVLADADGRLLYSDFLT 240  
QY 244 FLDRDSSKRLFYVETROPRARLLTLAHLFLVAPOHNOSEATG-----STSG----- 292  
DB 241 FLDRDSSKRLFYVETROPRARLLTLAHLFLVAPOHNOSEATG-----STSG----- 292  
QY 293 ---QALFASNVKPGORYVGE--GGQQLPASVSVSLREASGAVAPLTAGCTILINR 347  
DB 299 LGPRALFASVRPGORYVVAERDGRRLPLPAVHVSITLSEBAAGAVAPLTAGCTILINR 358  
QY 348 VLASCAVIEHSHMAHAFRFLAOGILAL-----CP 381  
DB 359 VLASCAVIEHSHMAHAFRFLAOGILAL-----CP 381  
QY 382 DGAIPATAATTGIIHWSRLLYRIGSVNLDGDLAHLPLGAVAPAS 425  
DB 419 GAADAPGATAGIHWYSQLLYIGTWLDSALHPLGMAVKSS 462

RESULT 15  
US-08-748-591-9  
; Sequence 9, Application US/08748591

Patent No. 5759811  
GENERAL INFORMATION:  
APPLICANT: Epstein, Ervin  
APPLICANT: Hu, Zhilan  
APPLICANT: Bonifas, Jeanette  
TITLE OF INVENTION: Mutant Human Hedgehog Gene  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish and Richardson  
STREET: 2200 Sand Hill Road  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/748,591  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J  
REGISTRATION NUMBER: 36,677  
REFERENCE/DOCKET NUMBER: 06510/067001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 322-5070  
TELEFAX: (415) 854-0875  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-748-591-9

Query Match 81.5%; Score 1807; DB 1; Length 462;  
Best Local Similarity 77.8%; Pred. No. 4,7e-195;  
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

QY 4 MLLTRLLVGFICALVSSGLTCGPGRGIGKRRHPKRLPLAYKOPIPVAEKTIGASG 63  
DB 1 MLLARCLLVLVSSLYCGLACGPGRGFGKRRHPKRLPLAYKOPIPVAEKTIGASG 60  
QY 64 RYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SYNWMPGV 123  
DB 61 RYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTORCKDKLNALAI SYNWMPGV 120  
QY 124 KLRTGEMDDGHHSESLHYEGRAVDITTSDDRKYGMILARLAVAGFDWYYESKAH 183  
DB 121 KLRTGEMDDGHHSESLHYEGRAVDITTSDDRKYGMILARLAVAGFDWYYESKAH 180  
QY 184 IHC5VKAENSVAAKSGCGPFGSATVHLEHGTRKLVKDLSPGDRVLADADGRLLYSDFLT 243  
DB 181 IHC5VKAENSVAAKSGCGPFGSATVHLEHGTRKLVKDLSPGDRVLADADGRLLYSDFLT 240  
QY 244 FLDRDSSKRLFYVETROPRARLLTLAHLFLVAPOHNOSEATG-----STSG----- 292  
DB 241 FLDRDSSKRLFYVETROPRARLLTLAHLFLVAPOHNOSEATG-----STSG----- 292  
QY 293 ---QALFASNVKPGORYVGE--GGQQLPASVSVSLREASGAVAPLTAGCTILINR 347  
DB 299 LGPRALFASVRPGORYVVAERDGRRLPLPAVHVSITLSEBAAGAVAPLTAGCTILINR 358  
QY 348 VLASCAVIEHSHMAHAFRFLAOGILAL-----CP 381  
DB 359 VLASCAVIEHSHMAHAFRFLAOGILAL-----CP 381  
QY 382 DGAIPATAATTGIIHWSRLLYRIGSVNLDGDLAHLPLGAVAPAS 425

Thu Feb 20 10:43:48 2003

us-09-827-110a-10.ra1

Page 10

DB 419 GAADAFGAGATAGIHWSOLYQIGTWLIDSEALHPGMAVKSS 462

Search completed: February 20, 2003, 10:14:10  
Job time : 13.1257 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:10:28 ; Search time 7.62902 Seconds  
(without alignments)  
1423.285 Million cell updates/sec

Title: US-09-827-110a-10

Sequence: 1 MEMLLRILRVGFCAL.....GSWLDGDLNPLGVAPAS 425

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 2554876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2218	100.0	425	US-08-900-220C-10	Sequence 10, Appl
2	2218	100.0	425	US-09-883-848A-10	Sequence 10, Appl
3	2218	100.0	425	US-09-021-660A-34	Sequence 34, Appl
4	2218	100.0	425	US-09-151-999-10	Sequence 10, Appl
5	1807	81.5	462	US-09-733-634-14	Sequence 14, Appl
6	1807	81.5	475	US-08-900-220C-15	Sequence 15, Appl
7	1807	81.5	475	US-09-883-848A-15	Sequence 15, Appl
8	1807	81.5	475	US-09-021-660A-39	Sequence 39, Appl
9	1807	81.5	475	US-09-151-999-15	Sequence 15, Appl
10	1791	80.7	437	US-09-990-046-14	Sequence 14, Appl
11	1791	80.7	437	US-09-969-520A-1	Sequence 1, Appl
12	1791	80.7	437	US-10-013-310-1	Sequence 1, Appl
13	1791	80.7	437	US-09-733-634-16	Sequence 16, Appl
14	1791	80.7	437	US-09-021-660A-37	Sequence 37, Appl
15	1788	80.6	437	US-08-900-220C-13	Sequence 13, Appl
16	1788	80.6	437	US-09-883-848A-13	Sequence 13, Appl
17	1788	80.6	437	US-09-151-999-13	Sequence 13, Appl
18	1785	80.5	437	US-09-969-520A-3	Sequence 3, Appl
19	1782	80.3	437	US-09-969-520A-10	Sequence 10, Appl

20	1782	80.3	437	9	US-09-969-520A-11	Sequence 11, Appl
21	1781	80.3	437	9	US-09-969-520A-2	Sequence 2, Appl
22	1777	80.1	437	9	US-09-969-520A-9	Sequence 9, Appl
23	1775	80.0	437	9	US-09-969-520A-4	Sequence 4, Appl
24	1771	79.8	437	9	US-09-969-520A-8	Sequence 8, Appl
25	1768	79.7	437	9	US-09-969-520A-6	Sequence 6, Appl
26	1759	79.3	437	9	US-09-969-520A-7	Sequence 7, Appl
27	1742	78.5	437	9	US-09-969-520A-5	Sequence 5, Appl
28	1494	67.4	418	10	US-09-021-660A-38	Sequence 38, Appl
29	1458	65.7	418	8	US-08-900-220C-18	Sequence 18, Appl
30	1458	65.7	416	9	US-09-883-848A-18	Sequence 18, Appl
31	1458	65.7	416	10	US-09-151-999-18	Sequence 18, Appl
32	1430	64.5	418	8	US-08-900-220C-14	Sequence 14, Appl
33	1430	64.5	418	9	US-09-883-848A-14	Sequence 14, Appl
34	1430	64.5	418	10	US-09-151-999-14	Sequence 14, Appl
35	1277.5	57.6	449	9	US-09-990-046-29	Sequence 29, Appl
36	1277.5	57.6	449	9	US-09-733-634-30	Sequence 30, Appl
37	1270.5	57.3	411	8	US-08-900-220C-16	Sequence 16, Appl
38	1270.5	57.3	411	9	US-09-883-848A-16	Sequence 16, Appl
39	1270.5	57.3	411	10	US-09-151-999-16	Sequence 16, Appl
40	1265.5	57.1	411	9	US-08-900-220C-12	Sequence 12, Appl
41	1249.5	56.3	411	8	US-08-900-220C-12	Sequence 12, Appl
42	1249.5	56.3	411	9	US-09-883-848A-12	Sequence 12, Appl
43	1249.5	56.3	411	10	US-09-151-999-12	Sequence 12, Appl
44	1200	54.1	396	8	US-08-900-220C-11	Sequence 11, Appl
45	1200	54.1	396	9	US-09-990-046-13	Sequence 13, Appl

#### ALIGNMENTS

RESULT 1  
US-08-900-220C-10  
Sequence 10, Application US/08900220C  
Patent No. US20020045206A1  
GENERAL INFORMATION:  
APPLICANT: Miao, Ningning  
Wang, Monica  
Mahantappa, Nagesh K.  
Jin, Ping  
Pang, Kevin  
TITLE OF INVENTION: Method of Treating Dopaminergic and GABA-nergic Disorders  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELLIOT LLP  
STREET: ONE POST OFFICE SQUARE  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII (text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/900,220C  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: ONV-044.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 425 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-08-900-220C-10

Query Match 100.0%; Score 2218; DB 8; Length 425;  
Best Local Similarity 100.0%; Pred. No. 7.8e-196;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGICGRHRPKLTPLAYKQFIPNVAEKTIG 60  
DB 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGICGRHRPKLTPLAYKQFIPNVAEKTIG 60  
OY 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMQW 120  
DB 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMQW 120  
OY 121 PGVKLRVTEGMDGDHSEESLHYEGRAVDITTSDRDSKYGMRLAVEAGFDWVYES 180  
DB 121 PGVKLRVTEGMDGDHSEESLHYEGRAVDITTSDRDSKYGMRLAVEAGFDWVYES 180  
OY 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRLLYS 240  
DB 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRLLYS 240  
OY 241 FLTFIDRMDSSRKLFLYVETROPARLLTAAHLFVAPQHNOSBASTSGQALFASNV 300  
DB 241 FLTFIDRMDSSRKLFLYVETROPARLLTAAHLFVAPQHNOSBASTSGQALFASNV 300  
OY 301 KPGQRYVYLGEQGQQLLPASVHVSRLREASGAVAPLTAOGTILINRVLASCYAVIEHS 360  
DB 301 KPGQRYVYLGEQGQQLLPASVHVSRLREASGAVAPLTAOGTILINRVLASCYAVIEHS 360  
OY 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTTGTHWYRLLYRIGSVLGDALHPGLM 420  
DB 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTTGTHWYRLLYRIGSVLGDALHPGLM 420  
OY 421 VAPAS 425  
DB 421 VAPAS 425

## RESULT 2

US-09-883-848A-10  
Sequence 10, Application US/09883848A  
Publication No. US20030022819A1  
GENERAL INFORMATION:  
APPLICANT: Ling, L.  
APPLICANT: Sanicola-Nadel, M.  
TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES  
FILE REFERENCE: CIBT-P01-119  
CURRENT FILING DATE: US/09/883, 848A  
CURRENT FILING DATE: 2001-06-18  
PRIOR APPLICATION NUMBER: 60/211, 919  
PRIOR FILING DATE: 2000-06-16  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 10  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Gallus gallus  
US-09-883-848A-10

Query Match 100.0%; Score 2218; DB 9; Length 425;  
Best Local Similarity 100.0%; Pred. No. 7.8e-196;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGICGRHRPKLTPLAYKQFIPNVAEKTIG 60  
DB 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGICGRHRPKLTPLAYKQFIPNVAEKTIG 60  
OY 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMQW 120  
DB 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMQW 120

OY 121 PGVKLRVTEGMDGDHSEESLHYEGRAVDITTSDRDSKYGMRLAVEAGFDWVYES 180  
DB 121 PGVKLRVTEGMDGDHSEESLHYEGRAVDITTSDRDSKYGMRLAVEAGFDWVYES 180  
OY 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRLLYS 240  
DB 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRLLYS 240  
OY 241 FLTFIDRMDSSRKLFLYVETROPARLLTAAHLFVAPQHNOSBASTSGQALFASNV 300  
DB 241 FLTFIDRMDSSRKLFLYVETROPARLLTAAHLFVAPQHNOSBASTSGQALFASNV 300  
OY 301 KPGQRYVYLGEQGQQLLPASVHVSRLREASGAVAPLTAOGTILINRVLASCYAVIEHS 360  
DB 301 KPGQRYVYLGEQGQQLLPASVHVSRLREASGAVAPLTAOGTILINRVLASCYAVIEHS 360  
OY 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTTGTHWYRLLYRIGSVLGDALHPGLM 420  
DB 361 WAHMAFAPFRLAOGLLAALCPDGAIPTAATTTTGTHWYRLLYRIGSVLGDALHPGLM 420  
OY 421 VAPAS 425  
DB 421 VAPAS 425

## RESULT 3

US-09-021-660A-34  
Sequence 34, Application US/09021660A  
Patent No. US20010041668A1  
GENERAL INFORMATION:  
APPLICANT: Baron, M.  
APPLICANT: Belaussoff, M.  
TITLE OF INVENTION: METHODS FOR MODULATING HEMATOPOIESIS AND VASCULAR  
TITLE OF INVENTION: GROWTH  
FILE REFERENCE: HOIP-P01-060  
CURRENT FILING DATE: US/09/021, 660A  
CURRENT FILING DATE: 2001-08-27  
PRIOR APPLICATION NUMBER: 60/037, 513  
PRIOR FILING DATE: 1997-02-10  
PRIOR APPLICATION NUMBER: 60/049, 763  
PRIOR FILING DATE: 1997-06-16  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 34  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Gallus gallus  
US-09-021-660A-34

Query Match 100.0%; Score 2218; DB 10; Length 425;  
Best Local Similarity 100.0%; Pred. No. 7.8e-196;  
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGICGRHRPKLTPLAYKQFIPNVAEKTIG 60  
DB 1 WEMLLLRLLLVGFCALVSSGLTCGPGRGICGRHRPKLTPLAYKQFIPNVAEKTIG 60  
OY 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMQW 120  
DB 61 ASGREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMQW 120  
OY 121 PGVKLRVTEGMDGDHSEESLHYEGRAVDITTSDRDSKYGMRLAVEAGFDWVYES 180  
DB 121 PGVKLRVTEGMDGDHSEESLHYEGRAVDITTSDRDSKYGMRLAVEAGFDWVYES 180  
OY 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRLLYS 240  
DB 181 KAHICSVKAENSVAAKSGCGPESATVHLEHGTRKLVKDLSPGDRVLAADADGRLLYS 240  
OY 241 FLTFIDRMDSSRKLFLYVETROPARLLTAAHLFVAPQHNOSBASTSGQALFASNV 300  
DB 241 FLTFIDRMDSSRKLFLYVETROPARLLTAAHLFVAPQHNOSBASTSGQALFASNV 300





```

1 STATE: MA
2 COUNTRY: USA
3 ZIP: 02109
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Floppy disk
7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: Ascii (text)
10
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/900,220C
13 FILING DATE: 24-Jul-1997
14 CLASSIFICATION: <unknown>
15 ATTORNEY/AGENT INFORMATION:
16 NAME: Vincent, Matthew P.
17 REGISTRATION NUMBER: 36,709
18 REFERENCE/DOCKET NUMBER: ONV-044,01
19
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (617) 832-1000
22 TELEFAX: (617) 832-7000
23
24 INFORMATION FOR SEQ ID NO: 15:
25 SEQUENCE CHARACTERISTICS:
26 LENGTH: 475 amino acids
27 TYPE: amino acid
28 TOPOLOGY: linear
29 MOLECULE TYPE: protein
30 SEQUENCE DESCRIPTION: SEQ ID NO: 15:
31
32 US-08-900-220C-15

```

```

: SEQ ID NO 39
: LENGTH: 475
: TYPE: PRF
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (463)
: OTHER INFORMATION: Xaa-unknown amino acid
: OS=99-021-660A-39

```

Query Match	81.5%	Score 1807;	DB 10	Length 475;
Best Local Similarity	77.8%	Pred. No. 5.3e-158;		
Matches 361; Conservative	22;	Mismatches 37;	Indels 44;	Gaps 6;

```

QY      4 MLTLRLILVLFICALVSSGLTGPCGSGIGRRRRPKTLPLATQOFTPNVAKETLGAG 63
Db      1 MLLLRCLLVLVSSLVCSGLACGPGGFGRRRRPKTLPLATQOFTPNVAKETLGAG 60

QY      64 RYEGKITNSERFEKELTPNYNDIIFKDEENTGADRLMTORCKDKMLALISVMNQPCV 123
Db      61 RYEGKITNSERFEKELTPNYNDIIFKDEENTGADRLMTORCKDKMLALISVMNQPCV 120

QY      124 KLRLTEGDEGHHSEESLHYEGRAVDITTSRDBSKGMRLRLVEAGFDMVNYESKHN 183
Db      121 KLRLTEGDEGHHSEESLHYEGRAVDITTSRDBSKGMRLRLVEAGFDMVNYESKHN 180

QY      184 IHCYSKAVNSYAAKSGCGFPGSATVHLEHGCKLVKDLSPGDRVLAADQGRILYSDFLT 243
Db      181 IHCYSKAVNSYAAKSGCGFPGSATVHLEHGCKLVKDLSPGDRVLAADQGRILYSDFLT 240

QY      244 FLDRDSEKRLFYVYIETQPRARLLTLTAHLFLVAPQHNOSATG-----STSG----- 292
Db      241 FLDRDQKAKKVFVYIETREPRERLLTLTAHLFLVAP--HND--ATEPEASSGSGPPSCGA 298

QY      293 ---QALFASNVKPGGRVYVLCG--GGQDLPLASVSVSLREESASAYAPLTQGTLLNR 347
Db      299 LGPRALFESRVRPGGRVYVVAERDQDRRLPLAAVHSVLSSEAAATYAPLTQGTILLNR 358

QY      348 VLASCVAVIEEHSMAHMAFAPRILAQGLLAL-----CP 381
Db      359 VLASCVAVIEEHSMAHMAFAPRILAHALLAALAPARTDGGDGGGDRGGCGGVALTAP 418

QY      382 DGAIPATAATTTGTHWYSRLLYRTGSWVLDGALPRLGVAAPAS 425
Db      419 GAADAPGAGATAGTHWYSQLLTGYTITLIDSEALPRLGVAAPAS 462

```

```

RESULT 9
US-09-151-999-15
: Sequence 15, Application US/09151999
: Patent No. US20020151460A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Elizabeth
: TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE
: TITLE OF INVENTION: POLYPEPTIDES, AND FORMULATIONS AND USES RELATED THERETO
: FILE REFERENCE: ONV-031.02
: CURRENT APPLICATION NUMBER: US/09/151,999
: CURRENT FILING DATE: 1998-08-11
: EARLIER APPLICATION NUMBER: 08/955,552
: EARLIER FILING DATE: 1997-10-20
: NUMBER OF SEQ ID NOS: 28
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 15
: LENGTH: 475
: TYPE: PRT
: ORGANISM: Homo sapien Shh
: FEATURE:
: OTHER INFORMATION: Xaa at position 463 is any or unknown amino acid
: US-09-151-999-15

```

Query Match	81.5%	Score 1807	DB 10	Length 475
Best Local Similarity	77.8%	Pred. No. 5.3e-158		
Matches 361	Conservative 22	Mismatches 37	Indels 44	Gaps 6

```

QY      4 MLITRILLYVFICALIYVSSGLTGPFGSGICKRHRHPKLTPLAKOFIPNVAETLIGASG 63
Db      1 MLLIARCLLLVAVSSSLVCSLACGPGGFCRRRHPKLTPLAKOFIPNVAETLIGASG 60
QY      64 RYEGKITRNSERFEKLPNNYNDIIFKDEENTGADRLTORCKDKLNALATSVNMOWPGV 123
Db      61 RYEEKISRNSERFEKLPNNYNDIIFKDEENTGADRLTORCKDKLNALATSVNMOWPGV 120
QY      124 KLRTYEGMDEDEGHSEESLHYEGRAVDITTSDBRBSYGMGLARLAVEGPMWVYESGAH 187
Db      121 KLRTYEGMDEDEGHSEESLHYEGRAVDITTSDBRBSYGMGLARLAVEGPMWVYESGAH 186
QY      184 IHCYSKAVNSVAANSGCCGFCPSATVHLEHGSTKLVKDLSPGDRVLADADRLYSDFLT 243
Db      181 IHCYSKAVNSVAANSGCCGFCPSATVHLEHGSTKLVKDLSPGDRVLADADGRLYSDFLT 244
QY      244 FLDRMDSRKLFYIETROPARLLTRAHLLFVAPOHNOSEATG-----STSG----- 297
Db      241 FLDRDGDGKKFYIETREPERRLLTRAHLLFVAHNS-ANGPEASSGSGSPPSGGA 296
QY      293 ---DALFASNVKPCORVYVLGE--GGQQLPASHVSYSLREASGAVAPLTAOGTILINR 347
Db      239 LGRPLRFLPSRRRPGORVYVAERBGDRRLTRAAVHASTVLTSEAGATAPLTAOGTILINR 355
QY      348 VLASCVAYIEEHSWANHMAFAPFRLAOGILLAL-----CP 381
Db      359 VLASCVAYIEEHSWANHMAFAPFRLAHALLALALAPARTDRGDSGGDRGGCGRVALTAP 411
QY      382 DGAIRPLATTTTGHWSRLLYRIGSWLDDDLHPRIGMAVPAS 425
Db      419 GAAADAPGAGATAGHWYSQLLTYOIGTWLTDSDEALHPRIGMAVKS 462

```

```

RESULT 10
US-09-990-046-14
: Sequence 14, Application US/09990046
: Patent No. US20020156245A1
: GENERAL INFORMATION:
: APPLICANT: de Sauvage, Frederic
: APPLICANT: Carpenter, David A.
: TITLE OF INVENTION: Patched-2
: FILE REFERENCE: P140581
: CURRENT APPLICATION NUMBER: US/09/990, 046
: CURRENT FILING DATE: 2001-11-20
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/293, 505
: PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-15
: NUMBER OF SEQ ID NOS: 32
: SEQ ID NO 14
: LENGTH: 437
: TYPE: PRT
: ORGANISM: Mus musculus
: US-09-990-046-14

```

	Query Match	80.7%	Score 1791	DB 9	Length 437
	Best Local Similarity	80.5%	Pred. No. 1.4e-156		
	Matches 354	Conservative 21	Mismatches 43	Indels 22	Gaps 6
QY	4	MLLLTRILLVGFICALVSSGLTGPCGNGIGKRRHPKLTLPPLAYKOFIPNVAEKTIGASG	63		
	1				
Db	2	LLLLLRCFLVLIASSLLVLCPCPGIACGPGGFGFRHRHPKLTLPPLAYKOFIPNVAEKTIGASG	61		
QY	64	RVEGKTIITNSERFPKLTIPNYNPDIIFKKEENCGARLTQRCRDKLNLALISVNNQMGCV	123		
	1				
Db	62	RVEGKTIITNSERFPKLTIPNYNPDIIFKKEENCGARLTQRCRDKLNLALISVNNQMGCV	121		
QY	124	KLRLVTEGWDDEGHHSEESLIHYEGRAVDITTTSDRDRSKYGLMARLRLAVEAGFDVYYEESKAH	183		
	1				
Db	122	KLRLVTEGWDDEGHHSEESLIHYEGRAVDITTTSDRDRSKYGLMARLRLAVEAGFDVYYEESKAH	181		
QY	184	THCSYKAANSVAANASGGCGPGSATYVHLEHGCTKLVKDLSPGDRVLAADADGRRLYSDFLT	243		
	1				
Db	182	THCSYKAANSVAANASGGCGPGSATYVHLEHGCTKLVKDLSPGDRVLAADADGRRLYSDFLT	241		

OY 244 FLDRDSSKLEFYIETROPARLLTLTAHLLFVAPQHNSSEATGSG-QALFASNVKP 302  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 242 FLDRDEGAKKVFYIETLEPRERLLTLTAHLLFVAP-HNDS---GPTGPSALFASRVAP 297  
OY 303 GORVYVLGE--GGQOLLPASVHSVSLREBASGAVAPLTAOQITILINRVLASCYAVIEHS 360  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 298 GORVYVVAERGGDRLLPAAVHSVTLREBEGAVAPLTAHGTILINRVLASCYAVIEHS 357  
OY 361 WAHMAFAPRLAOGLLALCP-----DGAIPTA-----ATTTGIMHYSRLYRI 405  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 358 WAHRAFAFAPRLAHALLAALAPARTDGGGGGSIIPAQOSATEARGAEPRTAGIMHYSOLLVHI 417  
OY 406 GSWVLDGALHPLGMAVAPAS 425  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
DB 418 GTWLIDSETHMPLGMAVKAS 437

RESULT 11  
US-09-969-520A-1  
; Sequence 1, Application US/09969520A  
; Patent No. US20020177163A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: BEACH, PHILIP A.  
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED  
; FILE REFERENCE: JH01670-1  
; CURRENT APPLICATION NUMBER: US/09/969,520A  
; PRIOR FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 60/235,153  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-969-520A-1

Query Match 80.7%; Score 1791; DB 9; Length 437;  
Best Local Similarity 80.5%; Pred. No. 1.4e-156;  
Matches 354; Conservative 21; Mismatches 43; Indels 22; Gaps 6;

OY 4 MILLTRILVGFICALVSSGLTGPGRGIGRRHPPKLTPLAYKQFIIPNAEKTIGASG 63  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
DB 2 LILLARCFVLIASSLLVCPGLACGPGRGFGRRHPPKLTPLAYKQFIIPNAEKTIGASG 61  
OY 64 RYEGKITRNSERFKELTPYNDIIFKDEBNTGADRLMTORCKDKLNLALTSVNMQPGV 123  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 62 RYEGKITRNSERFKELTPYNDIIFKDEBNTGADRLMTORCKDKLNLALTSVNMQPGV 121  
OY 124 KLRTVEGMEDEHNSHEESLHYEGRAVDITTSDRDSKYGMRLARLAVEGFDWVYIESKAH 183  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 122 KLRTVEGMEDEHNSHEESLHYEGRAVDITTSDRDSKYGMRLARLAVEGFDWVYIESKAH 181  
OY 184 IHCSTKAENSVAAGCGCPGSAIVHLEHGKTKLVKDLSPGDRVLAADQGHLYSDFLT 243  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 182 IHCSTKAENSVAAGCGCPGSAIVHLEHGKTKLVKDLSPGDRVLAADQGHLYSDFLT 241  
OY 244 FLDRDSSKLEFYIETROPARLLTLTAHLLFVAPQHNSSEATGSG-QALFASNVKP 302  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 242 FLDRDEGAKKVFYIETLEPRERLLTLTAHLLFVAP-HNDS---GPTGPSALFASRVAP 297  
OY 303 GORVYVLGE--GGQOLLPASVHSVSLREBASGAVAPLTAOQITILINRVLASCYAVIEHS 360  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 298 GORVYVVAERGGDRLLPAAVHSVTLREBEGAVAPLTAHGTILINRVLASCYAVIEHS 357  
OY 361 WAHMAFAPRLAOGLLALCP-----DGAIPTA-----ATTTGIMHYSRLYRI 405  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 358 WAHRAFAFAPRLAHALLAALAPARTDGGGGGSIIPAQOSATEARGAEPRTAGIMHYSOLLVHI 417  
OY 406 GSWVLDGALHPLGMAVAPAS 425  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

DB 418 GTWLIDSETHMPLGMAVKSS 437

RESULT 12  
US-10-013-310-1  
; Sequence 1, Application US/10013310  
; Publication No. US20020192216A1  
; GENERAL INFORMATION:  
; APPLICANT: Lamb, Jonathon Robert  
; APPLICANT: Hoyne, Gerard Francis  
; APPLICANT: Dallman, Margaret Jane  
; TITLE OF INVENTION: Therapeutic Use  
; FILE REFERENCE: 674525-2003  
; CURRENT APPLICATION NUMBER: US/10/013,310  
; PRIOR FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: PCT/GB00/02191  
; PRIOR FILING DATE: 2000-06-05  
; PRIOR APPLICATION NUMBER: UK 9913350.6  
; PRIOR FILING DATE: 1999-06-08  
; PRIOR APPLICATION NUMBER: UK 9921953.7  
; PRIOR FILING DATE: 1999-09-16  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: House Mouse  
US-10-013-310-1

Query Match 80.7%; Score 1791; DB 9; Length 437;  
Best Local Similarity 80.5%; Pred. No. 1.4e-156;  
Matches 354; Conservative 21; Mismatches 43; Indels 22; Gaps 6;

OY 4 MILLTRILVGFICALVSSGLTGPGRGIGRRHPPKLTPLAYKQFIIPNAEKTIGASG 63  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
DB 2 LILLARCFVLIASSLLVCPGLACGPGRGFGRRHPPKLTPLAYKQFIIPNAEKTIGASG 61  
OY 64 RYEGKITRNSERFKELTPYNDIIFKDEBNTGADRLMTORCKDKLNLALTSVNMQPGV 123  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 62 RYEGKITRNSERFKELTPYNDIIFKDEBNTGADRLMTORCKDKLNLALTSVNMQPGV 121  
OY 124 KLRTVEGMEDEHNSHEESLHYEGRAVDITTSDRDSKYGMRLARLAVEGFDWVYIESKAH 183  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 122 KLRTVEGMEDEHNSHEESLHYEGRAVDITTSDRDSKYGMRLARLAVEGFDWVYIESKAH 181  
OY 184 IHCSTKAENSVAAGCGCPGSAIVHLEHGKTKLVKDLSPGDRVLAADQGHLYSDFLT 243  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 182 IHCSTKAENSVAAGCGCPGSAIVHLEHGKTKLVKDLSPGDRVLAADQGHLYSDFLT 241  
OY 244 FLDRDSSKLEFYIETROPARLLTLTAHLLFVAPQHNSSEATGSG-QALFASNVKP 302  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 242 FLDRDEGAKKVFYIETLEPRERLLTLTAHLLFVAP-HNDS---GPTGPSALFASRVAP 297  
OY 303 GORVYVLGE--GGQOLLPASVHSVSLREBASGAVAPLTAOQITILINRVLASCYAVIEHS 360  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 298 GORVYVVAERGGDRLLPAAVHSVTLREBEGAVAPLTAHGTILINRVLASCYAVIEHS 357  
OY 361 WAHMAFAPRLAOGLLALCP-----DGAIPTA-----ATTTGIMHYSRLYRI 405  
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |  
DB 358 WAHRAFAFAPRLAHALLAALAPARTDGGGGGSIIPAQOSATEARGAEPRTAGIMHYSOLLVHI 417  
OY 406 GSWVLDGALHPLGMAVAPAS 425  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |  
DB 418 GTWLIDSETHMPLGMAVKSS 437

RESULT 13  
US-09-733-634-16  
; Sequence 16, Application US/09733634  
; Publication No. US20030013646A1  
; GENERAL INFORMATION:  
; APPLICANT: Massachusetts General Hospital  
; TITLE OF INVENTION: Method to stimulate insulin production by pancreatic b-cells

```

? FILE REFERENCE: 17653/1240
? CURRENT APPLICATION NUMBER: US/09/733,634
? PRIORITY FILING DATE: 2000-12-08
? PRIOR APPLICATION NUMBER: US 60/170,282
? PRIOR FILING DATE: 1995-12-10
? NUMBER OF SEQ ID NOS: 32
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 16
? LENGTH: 437
? TYPE: PRT
? ORGANISM: Mus musculus
US-09-733-634-16

```

Query Match	Score	DB	Length
80.78;	1791;	9;	437;

Oy	4	MLLRILLVIGFALCVSSGLTCGGRRIGRRHPRKLTPLAYKOFIRNVAEKTIGAGS	63
Db	2	LLLLARCFYLIIASSLIVPGIACGPRGFRGRHPRKLTPLAYKOFIRNVAEKTIGAGS	61
Oy	64	RYEKGITRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLINALAISVMNQPCV	123
Db	62	RYEKGITRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLINALAISVMNQPCV	121
Oy	124	KLRVTEGMDDECHHSEESIHYEGRAVDITTSRDRSKYCMRLARLVAEAFDWWYYSKRN	183
Db	122	KLRVTEGMDDECHHSEESIHYEGRAVDITTSRDRSKYCMRLARLVAEAFDWWYYSKRN	181
Oy	184	IHCISYKAENSVAAKSGGCGPGSATYHLEHGTKLYKDSPGDVRVLAADQGRLLYSDFLT	243
Db	182	IHCISYKAENSVAAKSGGCGPGSATYHLEHGTKLYKDSRGRVLAADQGRLLYSDFLT	241
Oy	244	FLDRDSSKRLFYVETROPARLLITAAHLLFVAPQHNOSEATGSG-OALFASNVKP	302
Db	242	FLDRDSSKRLFYVETROPARLLITAAHLLFVAPQHNOSEATGSG-OALFASNVKP	297
Oy	303	GGRVYVLAE--GGQDLPLASVHSVSLREBSAGAVAPLTQGTILIRVLAASCAVATEEHS	360
Db	298	GGRVYVLAE--GGQDLPLASVHSVSLREBSAGAVAPLTQGTILIRVLAASCAVATEEHS	357
Oy	361	WAHMAFAPRLAOGLLALCP-----DGAIPTA-----ATTGTGIHWYSRLLYRI	405
Db	358	WAHMAFAPRLAOGLLALCP-----DGAIPTA-----ATTGTGIHWYSRLLYRI	417
Oy	406	GSVVLDDGALHPLGVAPAS	425
Db	418	GSVVLDDGALHPLGVAPAS	437

```

1 RESULT 14
2 US-09-021-660A-37
3 : Sequence 37, Application US/09021660A
4 : Patent No. US20010041668A1
5 :
6 : GENERAL INFORMATION:
7 :
8 : APPLICANT: Baron, M.
9 : APPLICANT: Farrington, S.
10 : APPLICANT: Belaussoff, M.
11 : TITLE OF INVENTION: METHODS FOR MODULATING HEMATOPOIESIS AND VASCULAR
12 : TITLE OF INVENTION: GROWTH
13 : FILE REFERENCE: HUIP-P01-060
14 :
15 : CURRENT APPLICATION NUMBER: US/09/021,660A
16 : CURRENT FILING DATE: 2001-08-27
17 :
18 : PRIOR APPLICATION NUMBER: 60/037,513
19 : PRIOR FILING DATE: 1997-02-10
20 :
21 : PRIOR APPLICATION NUMBER: 60/049,763
22 : PRIOR FILING DATE: 1997-06-16
23 :
24 : NUMBER OF SEQ ID NOS: 42
25 :
26 : SOFTWARE: PatentIn Ver. 2.1
27 :
28 : SEQ ID NO 37
29 :
30 : LENGTH: 437
31 :
32 : TYPE: PRT
33 :
34 : ORGANISM: Mus musculus

```

US-09-021-660A-37

Query Match	80.7%;	Score 1791;	DB 10;	Length 437;
Best Local Similarity	80.5%;	Pred. No. 1.4e-156;		
Matches 354;	Conservative 21;	Mismatches 43;	Indels 22;	Gaps 6

```

Oy      4 MLLRLILVAFICALLVSLSLTCGPBGIGKRHHPKKLPLAYKOFIPVAEKTGGASC   63
         : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2 LLLLARCCLVLTLASSLVCPBACGPGRGCKRRHPPKKLPPLAYKOFIPVAEKTLGASG   61

Oy      64 RYEGKITRNSERFELPKNPNNDIIFKDEENTGADRLMTOCKDKMLAALISVMNOMPVG   1233
Db      62 RYEGKITRNSERFELPKNPNNDIIFKDEENTGADRLMTQRCOKDLAALISVMNOMPVG   1221

Oy      124 KLRTTECHDEEGHHSSESLHTEGAAVNDDITTSDBDRSKTYGMALARLAYAGCDWYYESKAH   1833
Db      122 KLRVTGEDEGHHSSESLHTEGAAVNDDITTSDBDRSKTYGMALARLAYAGCDWYYESKAH   1821

Oy      184 IHCSTKAENSVAAKSNGCGFCPSATAVNHLEHGCTKYKLOSGODRYLADADAGRLLYSDFLT   2433
Db      182 IHCSVKAKNSVAAKSNGCGFCPSATAVNHLEOGCTKYKLOPRGDRYLADADAGRLLYSDFLT   2411

Oy      244 FLDMDSRSKLFVIETROPARLLTAAHLLEFAPOHNOSEATGSTSG--QALFASNVPK    3022
Db      242 FLDRREGCAKFVYIETLEPERELLTAANHLLEFAVAP--HNDS---GTPGPSALFASFVR    2977

Oy      303 GORVVYVIGE--GGOOOLPVASHVSUSLREESAAGAAPTACTGILLNVLASCAVAIVEHS   3603
Db      298 GORYVYVAERGDRLRLPAAHVASVTLREEBEGAVALPARLAHGTILLNVLASCAVAIVEHS   3577

Oy      351 WAHMAFAEFRLAQGLIALCP-----DGAIPTA-----ATTTGTHMYSRLLYRI    4059
Db      358 WAHRAFAEFRLAHAILAALAAPARTDGGGGSIPAAGSATEARGAEPTAGIHMYSQLLYHI   4177

Oy      406 GSNVYLDGALAPLGVMNPAS   425
Db      418 GTWLIDSETMHPRLGNVAKSS   437

```

RESULT 15  
 US-08-900-220C-13  
 Sequence 13, Application US/08900220C  
 Patent No. US2002045206A1  
 GENERAL INFORMATION:  
 APPLICANT: Miao, Ningning  
 Wang, Monica  
 Mahanthappa, Nagesh K.  
 Pang, Kevin  
 Jin, Ping  
 TITLE OF INVENTION: Method of Treating Dopaminergic and  
 GABA-nergic Disorders  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
 STREET: ONE POST OFFICE SQUARE  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII (text)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/900,220C  
 FILING DATE: 24-Jul-1997  
 CLASSIFICATION: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Vincent, Matthew P.  
 REGISTRATION NUMBER: 36,709  
 REFERENCE/DOCKET NUMBER: ONV-044.01  
 TELECOMMUNICATION INFORMATION:



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:08:18 ; Search time 13.9665 Seconds  
(without alignments)  
2921.176 Million cell updates/sec

Title: US-09-827-110a-10  
Perfect score: 2218  
Sequence: 1 MVEMLLTRILLVGFICALL.....GSWLDGDLHPLHGVAPAS 425

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2218	100.0	425	2 A49424	patterning protein
2	1791	80.7	437	2 A49425	Sonic hedgehog pro
3	1780	80.3	437	2 B53193	hedgehog homolog v
4	1533	69.1	444	2 S56765	morphogen Xhh prec
5	1434	67.4	418	2 A53193	hedgehog homolog v
6	1429	64.4	415	2 A49426	sonic hedgehog gen
7	1200	54.1	396	2 B49425	Desert hedgehog pr
8	1052	47.4	336	2 C49425	Indian hedgehog pr
9	925	41.7	471	2 A46400	segment polarity p
10	449	20.2	94	2 G02735	desert hedgehog -
11	191.5	8.6	615	2 T29550	hypothetical prote
12	166	7.5	1207	2 T23754	hypothetical prote
13	158	7.1	1226	2 T24045	hypothetical prote
14	155.5	7.0	484	2 T34504	hypothetical prote
15	152.5	6.9	1021	2 T23252	hypothetical prote
16	142.5	6.4	481	2 T27665	hypothetical prote
17	130	5.9	629	2 T19563	hypothetical prote
18	128.5	5.8	868	2 T22281	hypothetical prote
19	115.5	5.2	313	2 T15855	hypothetical prote
20	114.5	5.2	481	2 T27975	hypothetical prote
21	105.5	4.8	1464	2 T13716	bazooka gene prote
22	103	4.6	205	2 T26220	hypothetical prote
23	102	4.6	477	2 G97611	UDP-N-acetyluramo
24	102	4.6	477	2 AC2834	UDP-MurNAC-pentape
25	102	4.6	503	2 C82068	aminopeptidase A/I
26	101	4.6	503	1 APECA	leucyl aminopeptid
27	101	4.6	503	2 E91283	aminopeptidase A/I
28	101	4.6	503	2 G86124	aminopeptidase A/I
29	100	4.5	503	2 AE1061	leucyl aminopeptid

30	99	4.5	811	2 F83451	probable cation-tr
31	98	4.4	1147	1 MMAXIB	myosin heavy chain
32	97.5	4.4	1137	2 JC5950	Integrin alpha-7 c
33	97	4.4	503	2 AC0418	leucyl aminopeptid
34	97	4.4	506	2 S55786	DNA nucleotidylexo
35	97	4.4	547	1 A23595	DNA nucleotidylexo
36	97	4.4	591	2 E83039	probable thiol-dis
37	96.5	4.4	1145	2 B75625	hypothetical prote
38	96	4.3	416	2 T02194	probable pectinase
39	95	4.3	893	2 E95053	cell wall surface
40	94.5	4.3	534	2 A32347	fimbrial protein t
41	94.5	4.3	598	2 AE2415	WD-repeat protein
42	94.5	4.3	924	2 T00518	hypothetical prote
43	94.5	4.3	1995	2 G81044	hemagglutinin/hemo
44	94.5	4.3	2124	2 T28658	polyketide synthas
45	94	4.2	481	2 G98044	UDP-N-acetyluramo

ALIGNMENTS

RESULT 1									
A49424									
patterning protein Sonic hedgehog precursor - chicken									
C:Species: Gallus gallus (chicken)									
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000									
C:Accession: A49424									
R:Riddle, R.D.; Johnson, R.L.; Laufer, E.; Tablin, C.									
A:Title: Sonic hedgehog mediates the polarizing activity of the ZPA.									
A:Reference number: A49424; MOID:94094333; PMID:8269518									
A:Accession: A49424									
A>Status: preliminary									
A:Molecule type: mRNA									
A:Residues: 1-425 <RID>									
A:Cross-references: GB:I28099; NID:9453526; PIDN:AAI72428.1; PID:9453527									
A:Superfamily: sonic hedgehog protein									
F:1-26/Domain: signal sequence #status predicted <SIG>									
QY	1	MVEMLLTRILLVGFICALVSSGLTGPCRGIGKRRHPRKPLPLAYKOPFVNAEFTLG	60						
DB	1	MVEMLLTRILLVGFICALVSSGLTGPCRGIGKRRHPRKPLPLAYKOPFVNAEFTLG	60						
QY	61	ASGRYBEKTRNSERFELTPNPNPDIIFKDEENTGADRLMTORCKDLNALAISVNNOW	120						
DB	61	ASGRYBEKTRNSERFELTPNPNPDIIFKDEENTGADRLMTORCKDLNALAISVNNOW	120						
QY	121	PGVLRVTEGMDDECHHSESLHYEGRAVDITTSDBRSRYGMARLAVENGFVWYVES	180						
DB	121	PGVLRVTEGMDDECHHSESLHYEGRAVDITTSDBRSRYGMARLAVENGFVWYVES	180						
QY	181	KAHHCYSKAKENSVAAKSGCFPGSATVHLHENGTKLVKDLSPDRVLADADGRLLYSD	240						
DB	181	KAHHCYSKAKENSVAAKSGCFPGSATVHLHENGTKLVKDLSPDRVLADADGRLLYSD	240						
QY	241	FLTFIDMDSSRKLFYIETROPARLLLTNAHLFTVAPOHNOSEATGTSGLAFASNV	300						
DB	241	FLTFIDMDSSRKLFYIETROPARLLLTNAHLFTVAPOHNOSEATGTSGLAFASNV	300						
QY	301	KPGGRVTVLEGGQQLPASHVSHVSLREKASGAAPLTAGCTILINNVLASCVVIEHS	360						
DB	301	KPGGRVTVLEGGQQLPASHVSHVSLREKASGAAPLTAGCTILINNVLASCVVIEHS	360						
QY	361	WAHAFAPFRIAGCLALCPDGAIPRAATTTGIIHYSRLLYRIGSWVLDGALHPLGM	420						
DB	361	WAHAFAPFRIAGCLALCPDGAIPRAATTTGIIHYSRLLYRIGSWVLDGALHPLGM	420						
QY	421	VAPAS 425							
DB	421	VAPAS 425							

Db 421 VAPAS 425

RESULT 2  
A:Accession: A49425  
C:Species: Sonic hedgehog protein precursor - mouse  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999  
R:Etchelard, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; McMe  
Cell 75, 1417-1430, 1993  
A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is implie  
A:Reference number: A49425; MUID:9409434; PMID:7916661  
A:Accession: A49425  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-437 <ECCH>  
A:Cross-references: GB:X76290  
A:Note: authors translation is shown for the codon TCC at position 436  
C:Genetics:  
A:Gene: Shh  
C:Superfamily: sonic hedgehog protein

Query Match 80.7%; Score 1791; DB 2; Length 437;  
Best Local Similarity 80.5%; Pred. No. 5,9e-141;  
Matches 354; Conservative 21; Mismatches 43; Indels 22; Gaps 6;

QY 4 MLTLRLLVGFICALVSSGLTCGPGRGIGRRHPPKLTPLAYKOFIPNVAEKTIGASG 63  
DB 2 LLLARCFVLIALSSLVCPGLACGPGRGGRHPPKLTPLAYKOFIPNVAEKTIGASG 61  
QY 64 RYEGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPGV 123  
DB 62 RYEGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPGV 121  
QY 124 KLRVTEGMDEDHSESLHYEGRAVDITTSDRDSKYGMRLAVALAEAGFDWVYYESKAH 183  
DB 122 KLRVTEGMDEDHSESLHYEGRAVDITTSDRDSKYGMRLAVALAEAGFDWVYYESKAH 181  
QY 184 IHCSYKAENSVAKSGCGPFSATVLEHGGTKLVKDLSPGDRVLAADOGRLYSDFLT 243  
DB 182 IHCSYKAENSVAKSGCGPFSATVLEHGGTKLVKDLSPGDRVLAADOGRLYSDFLT 241  
QY 244 FLDRDSSKRLFYVETROPARLLITAAHLFPAPQHNQSEATGSGQA-LFASNVKP 302  
DB 242 FLDRDSSKRLFYVETROPARLLITAAHLFPAPQHNQSEATGSGQA-LFASNVKP 297  
QY 303 GORVYVLEGE--GGQOLLPASVHSVLSREASGAYAPLTNOGITLNRVLAACYAIEHS 360  
DB 298 GORVYVLEGE--GGQOLLPASVHSVLSREASGAYAPLTNOGITLNRVLAACYAIEHS 357  
QY 361 MAHMAFAPRLAQLALACP-----DGAIPTA-----ATTGTGIMYSRLLYRI 405  
DB 358 MAHMAFAPRLAQLALACP-----DGAIPTA-----ATTGTGIMYSRLLYRI 403  
QY 406 GSWVLDGALHPLGMVAPAS 425  
DB 418 GTWLIDSETMPLGMVAPAS 437

RESULT 3  
B53193  
hedgehog homolog vhh-1 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999  
C:Accession: B53193  
R:Roelink, H.; Augburger, A.; Heemskerk, J.; Kozh, V.; Norlin, S.; Ruiz i Altaba, A.;  
Cell 76, 761-775, 1994  
A:Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedge  
A:Reference number: A53193; MUID:94170375; PMID:8124714  
A:Accession: B53193  
A:Status: preliminary  
A:Molecule type: mRNA

A:Residues: 1-437 <ROE>  
A:Cross-references: GB:L27340; NID:9452122; PIDN:AAA20999.1; PID:9452123  
C:Superfamily: sonic hedgehog protein

Query Match 80.3%; Score 1780; DB 2; Length 437;  
Best Local Similarity 80.0%; Pred. No. 4,9e-140;  
Matches 352; Conservative 22; Mismatches 44; Indels 22; Gaps 6;

QY 4 MLTLRLLVGFICALVSSGLTCGPGRGIGRRHPPKLTPLAYKOFIPNVAEKTIGASG 63  
DB 2 LLLARCFVLIALSSLVCPGLACGPGRGGRHPPKLTPLAYKOFIPNVAEKTIGASG 61  
QY 64 RYEGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPGV 123  
DB 62 RYEGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPGV 121  
QY 124 KLRVTEGMDEDHSESLHYEGRAVDITTSDRDSKYGMRLAVALAEAGFDWVYYESKAH 183  
DB 122 KLRVTEGMDEDHSESLHYEGRAVDITTSDRDSKYGMRLAVALAEAGFDWVYYESKAH 181  
QY 184 IHCSYKAENSVAKSGCGPFSATVLEHGGTKLVKDLSPGDRVLAADOGRLYSDFLT 243  
DB 182 IHCSYKAENSVAKSGCGPFSATVLEHGGTKLVKDLSPGDRVLAADOGRLYSDFLT 241  
QY 244 FLDRDSSKRLFYVETROPARLLITAAHLFPAPQHNQSEATGSGQA-LFASNVKP 302  
DB 242 FLDRDSSKRLFYVETROPARLLITAAHLFPAPQHNQSEATGSGQA-LFASNVKP 297  
QY 303 GORVYVLEGE--GGQOLLPASVHSVLSREASGAYAPLTNOGITLNRVLAACYAIEHS 360  
DB 298 GORVYVLEGE--GGQOLLPASVHSVLSREASGAYAPLTNOGITLNRVLAACYAIEHS 357  
QY 361 MAHMAFAPRLAQLALACP-----DGAIPTA-----ATTGTGIMYSRLLYRI 405  
DB 358 MAHMAFAPRLAQLALACP-----DGAIPTA-----ATTGTGIMYSRLLYRI 403  
QY 406 GSWVLDGALHPLGMVAPAS 425  
DB 418 GTWLIDSETMPLGMVAPAS 437

RESULT 4  
S56765  
morphogen xhh precursor - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 10-Oct-1995 #sequence\_revision 03-Jan-1995 #text\_change 21-Jul-2000  
C:Accession: S56765  
R:Stolow, M.A.; Shi, Y.B.  
Nucleic Acids Res. 23, 2555-2562, 1995  
A:Title: Xenopus sonic hedgehog as a potential morphogen during embryogenesis and thy  
A:Reference number: S56765; MUID:95357169; PMID:7630736  
A:Accession: S56765  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-444 <STO>  
A:Cross-references: EMBL:L39213; NID:9790937; PIDN:AAC42227.1; PID:9790938  
C:Superfamily: sonic hedgehog protein

Query Match 69.1%; Score 1533; DB 2; Length 444;  
Best Local Similarity 66.7%; Pred. No. 1.7e-119;  
Matches 301; Conservative 41; Mismatches 73; Indels 36; Gaps 5;

QY 4 MLTLRLLVGFICALVSSGLTCGPGRGIGRRHPPKLTPLAYKOFIPNVAEKTIGASG 62  
DB 1 MLVAQSLILLSFICTIVTPPLACGPGRGIGRRHPPKLTPLAYKOFIPNVAEKTIGASG 60  
QY 63 GRYEGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPGV 122  
DB 61 GRYEGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAIISVNMQPGV 120  
QY 123 VLRVTEGMDEDHSESLHYEGRAVDITTSDRDSKYGMRLAVALAEAGFDWVYYESKA 182  
DB 121 VLRVTEGMDEDHSESLHYEGRAVDITTSDRDSKYGMRLAVALAEAGFDWVYYESKA 180



OY 183 IHCSVKAENSVAAKSGCGPAGATVHLEHGCTKLKVDLSPEDRVLADADGRLLYSDFL 242  
 DB 181 IHCSVKAENSVAAKSGCGPAGATVHLEHGCTKLKVDLSPEDRVLADADGRLLYSDFL 240  
 OY 243 TFLDMDSSRKLFYVETFPQRRALLTFAHLLFVAPOHNSOEAAGSTSGALFASNVKP 302  
 DB 241 MTDDEKDVKKLFYVETFPQRRALLTFAHLLFVAPOHNSOEAAGSTSGALFASNVKP 294  
 OY 303 GORVVLLEGCGOQLLPASVSVSLREASGAVAPLTAOGTILINRYLASCYAVIEBESMA 362  
 DB 295 GDLITAPRKPTTLKAVKYEVDL-EEDTGATAPLTAHGTIVIDVYLASCYAVIEBHTMA 353  
 OY 363 HMAFAPRLAQLLAALCPDGAIPTAAT-----TTTG 394  
 DB 354 HMAFAPRLAQLLAALCPDGAIPTAAT-----TTTG 394  
 OY 395 IHWSRLLYRIGSWVLGDGALHPLGMVAPAS 425  
 DB 414 IHWSRLLYRIGSWVLGDGALHPLGMVAPAS 444

## RESULT 5

A53193

hedgehog homolog vhh-1 - zebra fish

C:Species: Brachydanio rerio (zebra fish)

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999

C:Accession: A53193

R:RogLink, H.; Augsburg, A.; Heemkerk, J.; Kozh, V.; Norlin, S.; Ruiz, I. Altabe, A.;

Cell 76, 761-775, 1994

A:Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgeh

A:Reference number: A53193; MUID:94170375; PMID:8124714

A:Accession: A53193

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-418 &lt;ROE&gt;

A:Cross-references: GB:127585; NID:9452159; P1D:9452160

C:Superfamily: sonic hedgehog protein

Query Match 67.4%; Score 1494; DB 2; Length 418;  
 Best Local Similarity 68.4%; Pred. No. 2.8e-116;

Matches 290; Conservative 50; Mismatches 76; Indels 8; Gaps 3;

OY 4 MLTLRILLVGFICALVSSGLTCGPGRCIGRRHPRKLTPLAYKQFIPNVAEKTLGASG 63  
 DB 1 MLTLRILLVLSLTLSLVSGIACGPGRCIGRRHPRKLTPLAYKQFIPNVAEKTLGASG 60  
 OY 64 RREGKTRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPCV 123  
 DB 61 RREGKTRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPCV 120  
 OY 124 KLRVTEGMDHGHSESLHYEGRAVDITTSRDNSKYGLMARLVAEAGFDWVYTESKAH 183  
 DB 121 KLRVTEGMDHGHSESLHYEGRAVDITTSRDNSKYGLMARLVAEAGFDWVYTESKAH 180  
 OY 184 IHCSVKAENSVAAKSGCGPAGATVHLEHGCTKLKVDLSPEDRVLADADGRLLYSDFL 243  
 DB 181 IHCSVKAENSVAAKSGCGPAGATVHLEHGCTKLKVDLSPEDRVLADADGRLLYSDFL 240  
 OY 244 FLDRDSSRKLFYVETFPQRRALLTFAHLLFVAPOHNSOEAAGSTSGALFASNVKP 303  
 DB 241 FLDRDSSRKLFYVETFPQRRALLTFAHLLFVAPOHNSOEAAGSTSGALFASNVKP 296  
 OY 304 QRVYVLLEGCGOQLLPASVSVSLREASGAVAPLTAOGTILINRYLASCYAVIEBESMAH 363  
 DB 297 QRVYVLLEGCGOQLLPASVSVSLREASGAVAPLTAOGTILINRYLASCYAVIEBESMAH 354  
 OY 364 WAFAPRLAQLLAALCPDGAIPTAAT--TTGIIHWSRLLYRIGSWVLGDGALHPLGMV 421  
 DB 355 WAFAPRLAQLLAALCPDGAIPTAAT--TTGIIHWSRLLYRIGSWVLGDGALHPLGMV 414  
 OY 422 APAS 425

DB 415 VNSS 418  
 RESULT 6  
 A49426  
 sonic hedgehog gene shh protein - zebra fish  
 N:Alternate names: local cell-cell interaction signaling protein  
 C:Species: Brachydanio rerio (zebra fish)  
 C:Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 10-Dec-1999  
 C:Accession: A49426  
 R:Krauss, S.; Concordet, J.P.; Ingham, P.W.  
 Cell 75, 1431-1444, 1993  
 A:Title: A functionally conserved homolog of the Drosophila segment polarity gene hh  
 A:Reference number: A49426; MUID:94094335; PMID:8269519  
 A:Accession: A49426  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-415 <KRA>  
 A:Experimental source: embryo  
 A:Note: sequence extracted from NCBI backbone (NCBIF:142459)  
 C:Superfamily: sonic hedgehog protein

Query Match 64.4%; Score 1429; DB 2; Length 415;  
 Best Local Similarity 66.4%; Pred. No. 7e-111;

Matches 283; Conservative 49; Mismatches 72; Indels 22; Gaps 5;

OY 4 MLTLRILLVGFICALVSSGLTCGPGRCIGRRHPRKLTPLAYKQFIPNVAEKTLGASG 63  
 DB 1 MLTLRILLVLSLTLSLVSGIACGPGRCIGRRHPRKLTPLAYKQFIPNVAEKTLGASG 60  
 OY 64 RREGKTRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPCV 123  
 DB 61 RREGKTRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQPCV 120  
 OY 124 KLRVTEGMDHGHSESLHYEGRAVDITTSRDNSKYGLMARLVAEAGFDWVYTESKAH 183  
 DB 121 KLRVTEGMDHGHSESLHYEGRAVDITTSRDNSKYGLMARLVAEAGFDWVYTESKAH 180  
 OY 184 IHCSVKAENSVAAKSGCGPAGATVHLEHGCTKLKVDLSPEDRVLADADGRLLYSDFL 243  
 DB 181 IHCSVKAENSVAAKSGCGPAGATVHLEHGCTKLKVDLSPEDRVLADADGRLLYSDFL 240  
 OY 244 FLDRDSSRKLFYVETFPQRRALLTFAHLLFVAPOHNSOEAAGSTSGALFASNVKP 303  
 DB 241 FLDRDSSRKLFYVETFPQRRALLTFAHLLFVAPOHNSOEAAGSTSGALFASNVKP 296  
 OY 304 QRVYVLLEGCGOQLLPASVSVSLREASGAVAPLTAOGTILINRYLASCYAVIEBESMAH 363  
 DB 297 QRVYVLLEGCGOQLLPASVSVSLREASGAVAPLTAOGTILINRYLASCYAVIEBESMAH 354  
 OY 364 WAFAPRLAQLLAALCPDGAIPTAATTTG--IHWYSR-----LVRIQSWVLGDGA 414  
 DB 355 WAFAPRLAQLLAALCPDGAIPTAATTTG--IHWYSR-----LVRIQSWVLGDGA 407  
 OY 415 LHPLGM 420  
 DB 408 LHPLGM 413  
 RESULT 7  
 B49425  
 Desert hedgehog protein precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999  
 C:Accession: B49425  
 R:Reichardt, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M  
 Cell 75, 1417-1430, 1993  
 A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is imp  
 A:Reference number: A49425; MUID:94094334; PMID:7916661  
 A:Accession: B49425  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-396 <ECH>



```

Oy 201 CFPSPATVHLEHGGRKLVKLDLSPGGRVLAADADRLTSDFLTELDPRDMSKKLFVYET 260
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 416 CFSGMEETE-DGKIMTKDKLIGDKVLSMD-EAFVYISPYIMFKRDEETAEFNLLET 473
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Oy 261 ROPRARLLTLTAHLFLFVAPOHNOSEATSGTSGOALFASNVKPGQRVYLVEGGQOLLRAS 320
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 474 ANGH8-ILTLDNHLIYVSDCRFRSDL-----KLVAAEYKMKDDCIHTTID-SNVYIKK 525
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Oy 321 VHSVLSREASGAVAPLTAQGTILINRVLASCYAVIEESHMAHMAFADPRLAOGLLAL- 379
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 526 VSKISKVIE-TGIYSPLESTGDIIVNRVLASCHSNMLAKSLQOTFESLYKRTSVFNHLM 584
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Oy 380 -----CPDGAIPATATTGTG 395
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 585 FPKSTEGDLPVGVETLTV 605
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

RESULT 12
T23754
hypothetical protein T05C12.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T23754; T24513
R:Thomas, K.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z19793
A:Accession: T23754
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1207 <MTL>
A:Cross-references: EMBL:Z49668; PIDN:CAA90265.1; GSPDB:GN00020; CESP:T05C12.10
A:Experimental source: clone M110
R:Burton, J.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z19901
A:Accession: T24513
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1207 <MT2>
A:Cross-references: EMBL:Z66500; PIDN:CAA91313.1; GSPDB:GN00020; CESP:T05C12.10
A:Experimental source: clone T05C12
C:Genetics:
A:Gene: CESP:T05C12.10
A:Map position: 2
A:introns: 31/3; 87/2; 141/3; 180/2; 203/3; 267/1; 776/2; 794/2; 834/2; 1086/3; 1143/3;
Query Match 7.5%; Score 166; DB 2; Length 1207;
Beat Local Similarity 29.3%; Pred. No. 3.4e-05;
Matches 55; Conservative 33; Mismatches 84; Indels 16; Gaps 6;

Oy 187 SVKAENSVAAASGC-----CFPGSATVHLEHGRTLVLDLSDGDRVLAADADGRLLYSDF 241
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 942 AVLADTPGAGAGGAGGRNSCSADSLV-TTVYGGKRMDELQIGDIVLVPSGNNLVKEYEV 1000
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Oy 242 LTFELDRMSSRKLFVIEIETROPARLLTLTAHLFLFVAPOHNOSEATSGTSG-----DAL 295
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1001 EMFYIKREKTRKNFVLLTKSGR-KLSLTGRLLFLVAESQVEQITMNPDG-IDVAMRESK 1059
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Oy 296 FASNVKPGQRVYLVEGGQOLLPASVHSVLSREASGAVAPLTAQGTILINRVLASCYAV 355
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1060 YAEKARKGECVLSIDESGEVLADEIVRVG--RMTNVGIYSPMTVGSLSLVDGVLSCFSH 1117
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Oy 356 IEESHMAH 363
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 1118 IESH8-AH 1124
      :||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

RESULT 13
T24045
hypothetical protein R08B4.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

```

C:Accession: T24045  
 R:White, S.  
 Submitted to the EMBL Data Library, November 1995  
 A:Reference number: Z19834  
 A:Accession: T24045  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1226 <WTL>  
 A:Cross-references: EMBL:Z68008; PIDN:CAA92000.1; GSPDB:GN00028; CESP:R08B4.1  
 A:Experimental source: clone R08B4  
 C:Genetics:  
 A:Gene: CESP:R08B4.1  
 A:Map position: X  
 A:Introns: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 839/3; 877/1; 960/3; 1

Query Match 7.1%; Score 158; DB 2; Length 1226;  
 Best Local Similarity 27.8%; Pred. No. 0.00016;  
 Matches 52; Conservative 37; Mismatches 82; Indels 16; Gaps 7;

OY 192 NSVAAKSGCGPFGSATVHLEHGKTLVNDLSPGDRVLAADADGRLYSDFLTDRMDS 251  
 DB 1012 SALVATGACFSLDTWV-TTPGKKRMDQIDIGYVLTADLE-KTYFPPTLMIHREPEK 1069  
 OY 252 KLFVIEROPRRALLTLTAHLF---VAPQHNSEATGSGALFASNVKPGQRYV 308  
 DB 1070 VOEFLTIMERGKT-LRITSRHMYRNKCGKSPQYIKMLPHDGAIRFASDLEVDCVVV 1128  
 OY 309 L--GCGGOOLLPASVHSLREASGAYAPLTAOCTILINRYLASCYAVIEHESWA---H 363  
 DB 1129 LYKGYRQOKIETITRSV-----RTGIYSPLNNRIIYNDLASCYSIQONTQTTF 1183  
 OY 364 WAFAPFR 370  
 DB 1184 WAYDKLR 1190

## RESULT 14

T34504  
 hypothetical protein ZK1290.12 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T34504

R:Atch, A.  
 Submitted to the EMBL Data Library, July 1995  
 A:Description: The sequence of C. elegans cosmid ZK1290.  
 A:Reference number: Z21535  
 A:Accession: T34504  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-484 <TAI>  
 A:Cross-references: EMBL:U21308; PIDN:AA893321.1; GSPDB:GN00020; CESP:ZK1290.12  
 A:Experimental source: strain Bristol N2; clone ZK1290  
 C:Genetics:  
 A:Gene: CESP:ZK1290.12  
 A:Map position: 2  
 A:Introns: 66/3; 124/2; 181/2; 392/1

Query Match 7.0%; Score 155.5; DB 2; Length 484;  
 Best Local Similarity 24.9%; Pred. No. 6.9e-05;  
 Matches 50; Conservative 50; Mismatches 82; Indels 19; Gaps 7;

OY 194 VAAKSGCGPFGSATVHLEHGKTLVNDLSPGDRVLAADADGRLYSDFLTDRMDS 253  
 DB 283 VASGVACFTGNSKY-WTPAGEKSMADSLVGDVMTYEV-GKMTYTRVASMILHRLPDKA 340  
 OY 254 LEVYETROPARRALLTLTAHLFVAPQHNSEATGSGALFASNVKPGQRYV 312  
 DB 341 AFIKLTED-----GAILDMPQHFYKANCVTSEMEVLAEDMTIGDCLAV--KE 389  
 OY 313 GOOLLPASVHSLREASGAYAPLTAOCTILINRYLASCYAVIE---EHSWAHAFAP 368  
 DB 390 NEKLVMTTISEKSTFE--TGVAAPWTEIGDLIVDVYASCHNVYKANTLSHTELFATSV 448

OY 369 FRLAOGLLAALCPDGAIPTPA 389  
 DB 449 OOKMRSVLGSLEBETHLPATS 469

## RESULT 15

T23252  
 hypothetical protein K02E2.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T23252

R:Lloyd, C.  
 Submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z19716  
 A:Accession: T23252  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1021 <WTL>  
 A:Cross-references: EMBL:Z81560; PIDN:CAB04547.1; GSPDB:GN00023; CESP:K02E2.2  
 A:Experimental source: clone K02E2  
 C:Genetics:  
 A:Gene: CESP:K02E2.2  
 A:Map position: 5  
 A:Introns: 41/1; 79/3; 131/3; 197/1; 262/1; 297/3; 566/3; 610/3; 735/3; 775/3; 834/3;

Query Match 6.9%; Score 152.5; DB 2; Length 1021;  
 Best Local Similarity 27.7%; Pred. No. 0.00036;  
 Matches 49; Conservative 31; Mismatches 72; Indels 25; Gaps 7;

OY 214 GTRKYKDLSPGDRVLAADADGRLYSDFLTDRMDSKLFYVETROPARRALLTLTAH 273  
 DB 814 GKRRDEIEIGYVLTADLK-TALFSATLWTHREPEVOEFLKIKTNGKT-LQLTAGH 871  
 OY 274 LLEFVA-----POHNOSEATGS-----TSGOALFASNVKPGQRYV 316  
 DB 872 FIYATECKRLPEKNSSLANSTPERYRLIDITLPDSEFKLASQLKIGECCLI--HNGDOF 929  
 OY 317 LPASVHSLREASGAYAPLTAOCTILINRYLASCYAVIEH---SWAHNAFAPFR 370  
 DB 930 RMQKIDSIS-KIVSTGIYSPLTENGRILVNDYLASCYSVQONVLTQTFEFNAFDRLR 985

Search completed: February 20, 2003, 10:13:30  
 Job time : 17.9865 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: February 20, 2003, 10:07:32 ; Search time 13.0329 Seconds

(without alignments)  
1352.533 Million cell updates/sec

Title: US-09-827-110a-10

Perfect score: 2218  
Sequence: 1 MVEMLLRRLVGFICALL.....GSWLDGDALHPLGVNAPAS 425

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2218	100.0	425	1 SHH_CHICK	091035 gallus gall
2	1807	81.5	462	1 SHH_HUMAN	015465 homo sapien
3	1791	80.7	437	1 SHH_MOUSE	062226 mus musculu
4	1780	80.3	437	1 SHH_RAT	063673 rattus norv
5	1603.5	72.3	432	1 SHH_CYNPY	090385 cynops pyr
6	1533	69.1	444	1 SHH_XENLA	092000 xenopus lae
7	1494	67.4	418	1 SHH_BRARE	092008 brachydantio
8	1458	65.7	416	1 TWHH_BRARE	090419 brachydantio
9	1295.5	58.4	409	1 IHH_CHICK	098938 gallus gall
10	1286	58.0	411	1 IHH_XENLA	091612 xenopus lae
11	1277.5	57.6	411	1 IHH_MOUSE	097812 mus musculu
12	1265.5	57.1	411	1 IHH_HUMAN	014673 homo sapien
13	1232.5	55.6	412	1 IHH_BRARE	098862 brachydantio
14	1200	54.1	396	1 DHH_MOUSE	061468 mus musculu
15	1198	54.0	396	1 DHH_HUMAN	043323 homo sapien
16	1129.5	50.9	366	1 DHH1_XENLA	091610 xenopus lae
17	1119.5	50.5	398	1 DHH2_XENLA	091611 xenopus lae
18	925	41.7	471	1 HH_DROME	002936 drosophila
19	875.5	39.5	481	1 HH_DROXY	056674 drosophila
20	603	27.2	121	1 SHH_RASEL	079828 rasbora ele
21	600	27.1	121	1 SHH_CARAU	079631 carassius a
22	600	27.1	121	1 SHH_PUNCO	079850 punctius tet
23	599	27.0	121	1 SHH_RASHE	079864 rasbora het
24	599	27.0	121	1 SHH_RASPA	079869 rasbora pav
25	598	27.0	121	1 SHH_TANAL	079915 tanichthys
26	597	26.9	121	1 SHH_AMBCH	079662 amblypharyn
27	595	26.8	121	1 SHH_DANVA	013235 danio aff.
28	595	26.8	121	1 SHH_DANVA	013235 danio aequi
29	595	26.8	121	1 SHH_DANAT	013238 danio aff.
30	595	26.8	121	1 SHH_DANFR	013245 danio frank
31	595	26.8	121	1 SHH_DANKE	079709 danio kerri
32	595	26.8	121	1 SHH_DANPU	079717 danio pulch
33	595	26.8	121	1 SHH_DEYDE	013241 devatio dev

34	595	26.8	121	1 SHH_DEYMA	013247 devatio mal
35	595	26.8	121	1 SHH_DEYPA	013250 devatio pat
36	591	26.6	121	1 SHH_PUNCO	079838 punctius con
37	396	17.9	88	1 DHH_BRARE	079729 brachydantio
38	293	13.2	58	1 IHH_CARAU	079663 carassius a
39	293	13.2	58	1 IHH_DANAT	013240 danio aff.
40	293	13.2	58	1 IHH_DANKE	079711 danio kerri
41	293	13.2	58	1 IHH_DANPU	079719 danio pulch
42	293	13.2	58	1 IHH_DEYDE	013243 devatio dev
43	293	13.2	58	1 IHH_PUNDE	079852 punctius tet
44	293	13.2	58	1 IHH_RASEL	079860 rasbora ele
45	289	13.0	58	1 SHH_PSEPR	079839 pseudorasbo

## ALIGNMENTS

RESULT 1	ID	SHH_CHICK	STANDARD;	PRT;	425 AA.
AC	091035;				
DT	15-JUL-1999 (Rel. 38, Created)				
DT	15-JUL-1999 (Rel. 38, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Sonic hedgehog protein precursor (SHH).				
GN	SHH.				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;				
OC	Gallus.				
ON	NCBI_TaxID=9031;				
OX	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Limb bud;				
RX	MEDLINE=94094333; PubMed=8269518;				
RA	Riddle R.D., Johnson R.L., Laufer E., Tablin C.;				
RT	"Sonic hedgehog mediates the polarizing activity of the ZPA.";				
RL	Cell 75:1401-1416(1993).				
RN	[2]				
RP	FUNCTION. AND AUTOPROTEOLYTIC CLEAVAGE.				
RX	MEDLINE=95254654; PubMed=7736596;				
RA	Koelink H., Porter J.A., Chiang C., Yanabe Y., Chang D.T.,				
RT	Beachy P.A., Jessell T.M.;				
RT	"Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of sonic hedgehog autoproteolysis.";				
RL	Cell 81:445-455(1995).				
CC	-1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REPRESSSES ANOTHER TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION.				
CC	-1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).				
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN THE POSTERIOR LIMB BUD MESENCHYME, THE HENSEN'S NODE, THE NOTOCHORD, AND THE FLOOR PLATE OF THE NEURAL TUBE.				
CC	-1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT STAGE 17 DURING THE INITIATION OF LIMB BUD FORMATION. FROM THAT POINT ONWARDS, THE EXPRESSION PATTERN EXACTLY MATCHES THE LOCATION OF THE ZONE OF POLARIZING ACTIVITY (ZPA).				
CC	-1- INDUCTION: BY RETINOIC ACID.				
CC	-1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN				

CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING. WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY.  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: L28099; AAA72428.1; -  
 DR HSP: 062226; 1VHH.  
 DR MEROPS: C46.001; -  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hntc.  
 DR InterPro: IPR003586; Hedgehog\_hntc.  
 DR InterPro: IPR003587; Hedgehog\_hntc.  
 DR InterPro: IPR002203; Interh.  
 DR Pfam: PF01079; Hntc.1.  
 DR Pfam: PF01085; HH\_signal.1.  
 DR PRINTS: PR00632; SONICHHOG.  
 DR PRODOM: PD003042; HH\_signal.1.  
 DR SMART: SM00305; Hntc.1.  
 DR SMART: SM00306; Hntc.1.  
 DR PROSITE: PS00817; INTER\_N\_TER.1.  
 DR Developmental protein: Autocatalytic cleavage; Hydrolase; Protease;  
 DR Signal: Lipoprotein; Palmitate.  
 FT SIGNAL 1 26  
 FT CHAIN 27 425  
 FT CHAIN 27 200  
 FT CHAIN 201 425  
 FT SITE 200 201  
 FT SITE 246 246  
 FT SITE 270 270  
 FT ACT\_SITE 273 273  
 FT BINDING 200 200  
 FT DOMAIN 390 393  
 FT LIPID 27 27  
 SQ SEQUENCE 425 AA: 46474 MW: 49627.443040173 CRC64:  
 Query Match 100.0%; Score 2218; DB 1; Length 425;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-172; Indels 0; Gaps 0;  
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MVELLLRILLVGFICALVSSGLTGFGRGIGRRHKKLLPLAYKQIFPNVAEKTIG 60  
 1 MVELLLRILLVGFICALVSSGLTGFGRGIGRRHKKLLPLAYKQIFPNVAEKTIG 60  
 1 MVELLLRILLVGFICALVSSGLTGFGRGIGRRHKKLLPLAYKQIFPNVAEKTIG 60  
 61 ASGRGKTRRSEKPELTPNNDITRKDEENGCADLMQRCKDKNALAIVSNQW 120  
 61 ASGRGKTRRSEKPELTPNNDITRKDEENGCADLMQRCKDKNALAIVSNQW 120  
 61 ASGRGKTRRSEKPELTPNNDITRKDEENGCADLMQRCKDKNALAIVSNQW 120  
 121 PGVKLRVTEGMEDGHHSESLHYEGRAVDITSDRDRSKYGLARLAVEAGFDWYYS 180  
 121 PGVKLRVTEGMEDGHHSESLHYEGRAVDITSDRDRSKYGLARLAVEAGFDWYYS 180  
 121 PGVKLRVTEGMEDGHHSESLHYEGRAVDITSDRDRSKYGLARLAVEAGFDWYYS 180  
 181 KAHHCSSVAENSVAAKSGCGPGSATVHLHGGTKLVKDLSPGGRVLAADADGRLLYS 240  
 181 KAHHCSSVAENSVAAKSGCGPGSATVHLHGGTKLVKDLSPGGRVLAADADGRLLYS 240  
 181 KAHHCSSVAENSVAAKSGCGPGSATVHLHGGTKLVKDLSPGGRVLAADADGRLLYS 240  
 241 FLTFDRMDSSKRLFYVETTRPARLLTLAHLFVAPQHNSERTGSTSQALFASNV 300  
 241 FLTFDRMDSSKRLFYVETTRPARLLTLAHLFVAPQHNSERTGSTSQALFASNV 300

DB 241 FLTFDRMDSSKRLFYVETTRPARLLTLAHLFVAPQHNSERTGSTSQALFASNV 300  
 QY 301 KGORVYVLEGEGGQDLIPASVSVSLREBASCAYAPLTAQGTILLNRVASCYAVIEHS 360  
 DB 301 KGORVYVLEGEGGQDLIPASVSVSLREBASCAYAPLTAQGTILLNRVASCYAVIEHS 360  
 QY 361 WAHMAFAPRLAOGLLAALCPDGAIPTAATTTTGGTGHWSRLLYRIGSVVLDGDLHPGLM 420  
 DB 361 WAHMAFAPRLAOGLLAALCPDGAIPTAATTTTGGTGHWSRLLYRIGSVVLDGDLHPGLM 420  
 QY 421 VAPAS 425  
 DB 421 VAPAS 425  
 RESULT 2  
 SHH\_HUMAN STANDARD; PRT; 462 AA.  
 ID SHH\_HUMAN  
 AC Q15465;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sonic hedgehog protein precursor (SHH) (HHC-1).  
 GN SHH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal Lung; PubMed=7590746;  
 RX MEDLINE=96070431; PubMed=7590746;  
 RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,  
 RA Gasier J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seidman C.E.,  
 RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.;  
 RT "Cloning, expression, and chromosomal location of SHH and HH: two  
 RT human homologues of the Drosophila segment polarity gene hedgehog";  
 RL Genomics 28:44-51(1995).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RA Tate G., Kishimoto K., Mitsuya T.;  
 RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in  
 RT human cancer cell lines and embryonic organs";  
 RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).  
 [3]  
 RP SEQUENCE OF 1-187 FROM N.A.  
 RA Strong C., Graves T., Sutterer C., Ozerisky P.;  
 RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.  
 [4]  
 RP SEQUENCE OF 119-167 FROM N.A.  
 RX MEDLINE=95236997; PubMed=7720571;  
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,  
 RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;  
 RT "Products, genetic linkage and limb patterning activity of a murine  
 RT hedgehog gene";  
 RL Development 120:3339-3353(1994).  
 [5]  
 RP PALMITOYLATION.  
 RX MEDLINE=98256339; PubMed=9593755;  
 RA Pepsinsky R.B., Zeng C., Wen D., Rayhorn P., Baker D.P., Williams K.P.,  
 RA Bixler S.A., Ambrose C.M., Garber E.A., Mlatkowski K., Taylor F.R.,  
 RA Wang E.A., Galdes A.;  
 RT "Identification of a palmitic acid-modified form of human Sonic  
 RT hedgehog";  
 RL J. Biol. Chem. 273:14037-14045(1998).  
 [6]  
 RP VARIANTS HPE ARG-31; GLY-117 AND ARG-117.  
 RX MEDLINE=97051937; PubMed=8896572;  
 RA Roessler E., Belloni E., Gaudent K., Jay P., Berta P., Scherer S.W.,  
 RA Tsui L.-C., Muenke M.;  
 RT "Mutations in the human Sonic hedgehog gene cause holoprosencephaly";  
 RL Nat. Genet. 14:357-360(1996).  
 [7]

RP VARIANT: HPE ARG-31; GLY-117; ARG-117; GLU-224; THR-226 AND THR-383.  
RX MEDLINE=98027056; PubMed=9302262;  
RA Roessler E., Belloni E., Gaudenz K., Vargas F., Scherer S.W.,  
RA Teul L.-C., Muenke M.;  
RT "Mutations in the C-terminal domain of Sonic hedgehog cause  
RT holoprosencephaly";  
RL Hum. Mol. Genet. 6:1847-1853(1997).  
RN [9]  
RP VARIANTS: HPE HIS-100; GLN-188 AND ASN-222.  
RX MEDLINE=99371775; PubMed=1041331;  
RA Odent S., Attli-Bitach T., Blayau M., Mathieu M., Aug J.,  
RA Delgado de A.L., Gall J.Y., Le Marec B., Munnich A., David V.,  
RA Vekemans M.;  
RT "Expression of the Sonic hedgehog (SHH) gene during early human  
RT development and phenotypic expression of new mutations causing  
RT holoprosencephaly";  
RL Hum. Mol. Genet. 8:1683-1689(1999).  
RN [9]  
RP VARIANTS: HPE V-88; K-115; R-236; 263-R--A-269 DEL; D-290; A-424 AND  
RP L-436.  
RX MEDLINE=20025757; PubMed=10556296;  
RA Nanni L., Ming J.E., Bocian M., Steinhans K., Bianchi D.M.,  
RA Die-Smulders C., Giannotti A., Imazumi K., Jones K.L., Campo M.D.,  
RA Martin R.A., Meinecke P., Plempont M.E.M., Robin N.H., Young I.D.,  
RA Roessler E., Muenke M.;  
RT "The mutational spectrum of the sonic hedgehog gene in  
RT holoprosencephaly: SHH mutations cause a significant proportion of  
RT autosomal dominant holoprosencephaly";  
RL Hum. Mol. Genet. 8:2479-2488(1999).  
CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN  
CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION  
CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE  
CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER  
CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A  
CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED  
CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE  
CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE  
CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. THRESHOLD  
CC FLOOR PLATE-AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD  
CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS  
CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY  
CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN FETAL INTESTINE, LIVER, LUNG, AND  
CC KIDNEY. NOT EXPRESSED IN ADULT TISSUES.  
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY  
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
CC -1- DISEASE: DEFECTS IN SHH ARE THE CAUSE OF HOLOPROSENCEPHALY (HPE).  
CC HPE IS A CLINICALLY VARIABLE AND GENETICALLY HETEROGENEOUS  
CC MALFORMATION IN WHICH THE DEVELOPING FOREBRAIN FAILS TO CORRECTLY  
CC SEPARATE INTO RIGHT AND LEFT HEMISPHERES. IN THE MOST SEVERE FORM  
CC (ALOUP HPE), THERE IS NO INTERHEMISPHERIC FISSURE, A SINGLE BRAIN  
CC VENTRICLE IS PRESENT. HPE IS ASSOCIATED WITH SEVERAL DISTINCT  
CC FACIES AND PHENOTYPIC VARIABILITY. IN THE MOST EXTREME CASES,  
CC ANOPHTHALMIA OR CYCLOPIA IS EVIDENT ALONG WITH A CONGENITAL  
CC ABSENCE OF THE MATURE NOSE. THE LESS SEVERE FORM FEATURES FACIAL  
CC DYSMORPHIA CHARACTERIZED BY OCULAR HYPERTELORISM, DEFECTS OF THE  
CC UPPER LIP AND/OR NOSE, AND ABSENCE OF THE OLFACTORY NERVES OR  
CC CORPUS CALLOSUM. THE MAJORITY OF HPE CASES ARE APPARENTLY  
CC SPORADIC, ALTHOUGH CLEAR EXAMPLES OF AUTOSOMAL DOMINANT (AD)  
CC INHERITANCE HAVE BEEN DESCRIBED. INTERESTINGLY, UP TO 30% OF  
CC OBLIGATE CARRIERS OF HPE GENE IN AD PEDIGREES ARE CLINICALLY  
CC UNAFFECTED.

Query Match	81.5%	Score 1807	DB 1	Length 462
CC	-1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.			
CC	-1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;			
CC	WWW="http://www.infobiosgen.fr/services/chromcancer/Genes/SHHID378.html"			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_licence@isb-sib.ch">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:licence@isb-sib.ch">licence@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: L36518; AAA62179.1; -.			
DR	EMBL: AC002484; AAB67604.1; -.			
DR	HSSP: Q62226; 1VHH.			
DR	MEROPS: C46_002; -.			
DR	Genew: HGNC:10848; SHH.			
DR	MIM: 600725; -.			
DR	MIM: 142945; -.			
DR	InterPro: IPR000320; HH_signal.			
DR	InterPro: IPR001767; Hedgehog_hntc.			
DR	InterPro: IPR003586; Hedgehog_hntc.			
DR	InterPro: IPR003587; Hedgehog_hntn.			
DR	InterPro: IPR002203; Intein.			
DR	InterPro: IPR001657; SonicHH.			
DR	Pfam: PF01079; Hntc; 1.			
DR	Pfam: PF01085; HH_signal; 1.			
DR	PRINTS: PR00632; SONICHHOG.			
DR	ProDom: PD003042; HH_signal; 1.			
DR	SMART: SM00305; Hntc; 1.			
DR	SMART: SM00306; Hntn; 1.			
DR	PROSITE: PS50817; INTEIN_N_TER; 1.			
KM	Developmental protein: Autocatalytic cleavage; Hydrolyase; Protease;			
KM	Signal: Lipoprotein; Palmitate; Disease mutation; Holoprosencephaly.			
FT	SIGNAL	1	23	POTENTIAL.
FT	CHAIN	24	462	SONIC HEDGEHOG PROTEIN.
FT	CHAIN	24	197	SONIC HEDGEHOG PROTEIN N-PRODUCT.
FT	CHAIN	198	462	SONIC HEDGEHOG PROTEIN C-PRODUCT.
FT	SITE	197	198	CLEAVAGE (AUTO-) (BY SIMILARITY).
FT	SITE	243	243	INVOLVED IN CHOLESTEROL TRANSFER (BY
FT	SITE	267	267	SIMILARITY).
FT	SITE	267	267	INVOLVED IN AUTO-CLEAVAGE (BY
FT	ACT_SITE	270	270	SIMILARITY).
FT	ACT_SITE	270	270	ESSENTIAL FOR AUTO-CLEAVAGE (BY
FT	BINDING	197	197	SIMILARITY).
FT	BINDING	407	411	CHOLESTEROL (BY SIMILARITY).
FT	DOMAIN	24	24	POLY-GLY.
FT	LIPID	31	31	CHOLESTEROL (BY SIMILARITY).
FT	VARIANT	88	88	PALMITATE.
FT	VARIANT	88	88	G -> R (IN HPE).
FT	VARIANT	100	100	D -> V (IN HPE: FAM11AL).
FT	VARIANT	100	100	/FTId-VAR_009163.
FT	VARIANT	115	115	Q -> H (IN HPE: SPORADIC).
FT	VARIANT	117	117	/FTId-VAR_009164.
FT	VARIANT	117	117	N -> R (IN HPE: FAM11AL).
FT	VARIANT	117	117	/FTId-VAR_009165.
FT	VARIANT	117	117	W -> G (IN HPE).
FT	VARIANT	188	188	/FTId-VAR_003620.
FT	VARIANT	188	188	W -> R (IN HPE).
FT	VARIANT	222	222	/FTId-VAR_003621.
FT	VARIANT	222	222	E -> Q (IN HPE: FAM11AL).
FT	VARIANT	222	222	/FTId-VAR_009166.
FT	VARIANT	224	224	D -> N (IN HPE: FAM11AL).
FT	VARIANT	224	224	/FTId-VAR_009167.
FT	VARIANT	226	226	V -> E (IN HPE).
FT	VARIANT	226	226	/FTId-VAR_009168.
FT	VARIANT	236	236	A -> T (IN HPE: FAM11AL).
FT	VARIANT	236	236	/FTId-VAR_009169.
FT	VARIANT	269	269	S -> R (IN HPE: FAM11AL).
FT	VARIANT	269	269	/FTId-VAR_009170.
FT	VARIANT	269	269	MISSING (IN HPE: SPORADIC).



Best Local Similarity 77.8%; Pred. No. 4,4e-139;  
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

QY 4 MLTTRILVGCIFICALIVSSGLTGPCRGIGKRRHKKLPKLPYAKOPIPVNAEKTLCASG 63  
1 MLTTRILVGCIFICALIVSSGLTGPCRGIGKRRHKKLPKLPYAKOPIPVNAEKTLCASG 60

QY 64 RYEGKTRNSEREREKELTPNTPDITFEDEENTGADRLMTCRKDKLALAI SVANQMPGV 123  
61 RYEGKTRNSEREREKELTPNTPDITFEDEENTGADRLMTCRKDKLALAI SVANQMPGV 120

QY 124 KLRTGEGMDGDHSESLYEGRAVDITSDRNRKYGMRLARVAGRDWYTESKAH 183  
121 KLRTGEGMDGDHSESLYEGRAVDITSDRNRKYGMRLARVAGRDWYTESKAH 180

QY 184 IHCYKAKNSVAAKSGGCGFSGATVHLEHSGTKLVKDLSPGDRYLAADAGRLYSDFLT 243  
181 IHCYKAKNSVAAKSGGCGFSGATVHLEHSGTKLVKDLSPGDRYLAADAGRLYSDFLT 240

QY 244 FLDRMDSKRLFYETFRORARILLTAHLLFYAPQHNSGATG-----STSG----- 292  
241 FLDRMDSKRLFYETFRORARILLTAHLLFYAPQHNSGATG-----STSG----- 298

QY 293 ---GALFASNVKPGORYVYGE--GGQQLLPASVHSVSLREASGAPAPLTAQSTILLNR 347  
299 LGRRALFASNVKPGORYVYGE--GGQQLLPASVHSVSLREASGAPAPLTAQSTILLNR 358

QY 348 VLASCVAVIEHSHVMAHAFAPFLAOGILAL-----CP 381  
359 VLASCVAVIEHSHVMAHAFAPFLAOGILAL-----CP 381

QY 382 DGAIPRTAATTTTGHWSKLLRYIGSVNVDGDLHPGLGAPVAPS 425  
419 DGAIPRTAATTTTGHWSKLLRYIGSVNVDGDLHPGLGAPVAPS 462

DB 419 GAADAPGAGATAGIHWSQLLYQIGTWLIDSEALHPGLGAPVAPS 462

RESULT 3  
SHH\_MOUSE STANDARD: PRT; 437 AA.

ID SHH\_MOUSE  
AC 062226;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Sonic hedgehog protein precursor (SHH) (HMG-1).  
GN SHH OR HHG1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Embryo;  
RX MEDLINE=94094334; PubMed=7916661;  
RA Echeveria Y., Epstein D.J., St Jacques B., Shen L., Mohler J., McMahon J.A., McMahon A.P.;  
RT "Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity."; Cell 75:1417-1430(1993).  
RL (2)  
RN REVISION TO 122.  
RC STRAIN=C57BL/6J;  
RA McMahon A.P.;  
RN Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
RN (3)  
RP SEQUENCE FROM N.A. AND AUTOPROTEOLYTIC CLEAVAGE.  
RX MEDLINE=95234654; PubMed=7720571;  
RA Chang D.T., Lopez A., von Kries D.P., Chiang C., Simandl B.K., Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;  
RT "Products, genetic linkage and limb patterning activity of a murine hedgehog gene."; Development 120:3339-3353(1994).  
RL (4)  
RP FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.

RX MEDLINE=95234654; PubMed=7736596;  
RA Roellink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T., Beachy P.A., Jessell T.M.;  
RT "Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of sonic hedgehog autoproteolysis."; Cell 81:445-455(1995).  
RL (5)  
RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 34-195.  
RP MEDLINE=96069744; PubMed=7477329;  
RA Hall T.M.T., Porter J.A., Beachy P.A., Leahy D.J.;  
RT "A potential catalytic site revealed by the 1.7-A crystal structure of the amino-terminal signalling domain of Sonic hedgehog."; Nature 378:212-216(1995).  
RL  
-1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH FLOOR PLATE-AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY SIMILARITY).  
-1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL. WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
-1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF EMBRYONIC TISSUES INCLUDING THE NOTOCHORD, VENTRAL NEURAL TUBE, FLOOR PLATE, LUNG BUD, ZONE OF POLARIZING ACTIVITY AND POSTERIOR DYSPL. MESENCHYME OF LIMBS. IN THE ADULT, EXPRESSED IN LUNG AND NEURAL RETINA.  
-1- DEVELOPMENTAL STAGE: FIRST DETECTABLE DURING GASTRULATION.  
-1- INDUCTION: BY RETINOIC ACID.  
-1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY.  
-1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sb.ch/announce/> or send an email to [license@isb-sb.ch](mailto:license@isb-sb.ch)).  
CC  
CC EMBL; X76290; CAA53922.1; -  
CC PDB; 1VHH; 29-JAN-96.  
CC MEROPS; C46.002; -  
CC MGD; MGI:98297; Shh.  
DR InterPro: IPR000320; HH\_signal.  
DR InterPro: IPR001767; Hedgehog\_hint.  
DR InterPro: IPR003586; Hedgehog\_hint.  
DR InterPro: IPR003587; Hedgehog\_hint.  
DR InterPro: IPR002203; Intein.  
DR InterPro: IPR001657; SotichH.  
DR Pfam; PF01079; Hint; 1.  
DR Pfam; PF01085; HH\_signal; 1.  
DR PRINTS; PRO0632; SONICHHOG.  
DR PRODOM; PD003042; HH\_signal; 1.  
DR SMART; SM00305; HintC; 1.  
DR SMART; SM00306; HintN; 1.  
DR PROSITE; PS50817; INTEIN\_N\_TER; 1.



KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KM Signal; Lipoprotein; Palmitate; 3D-structure.  
 FT CHAIN 1 24 POTENTIAL.  
 FT CHAIN 25 437 SONIC HEDGEHOG PROTEIN.  
 FT CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.  
 FT CHAIN 199 437 SONIC HEDGEHOG PROTEIN C-PRODUCT.  
 FT SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).  
 FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY  
 FT SITE 244 244 SIMILARITY).  
 FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY  
 FT ACT\_SITE 271 271 SIMILARITY).  
 FT BINDING 198 198 ESSENTIAL FOR AUTO-CLEAVAGE (BY  
 FT DOMAIN 383 387 CHOLESTEROL (BY SIMILARITY).  
 FT LIPID 25 25 POLY-GLY.  
 SQ SEQUENCE 437 AA; 47773 MM; DOEB72F08E7860DF CRC64;  
 Query Match 80.7%; Score 1791; DB 1; Length 437;  
 Best Local Similarity 80.5%; Pred. No. 8e-138;  
 Matches 354; Conservative 21; Mismatches 43; Indels 22; Gaps 6;

QY 4 MLTLRLVLFICALVSSGLTGGPGGICGRHRRPKLTPLAKQFIPNAEKTGLASG 63  
 DB 2 LLLLRGLVLIASSLVCPGLACGPGGFGRRHRRPKLTPLAKQFIPNAEKTGLASG 61  
 QY 64 RYEEKITRNSERFKEITPNYPNDIIFKDENTGADRLMTORCKDLNALAISVMNOMPGV 123  
 DB 62 RYEEKITRNSERFKEITPNYPNDIIFKDENTGADRLMTORCKDLNALAISVMNOMPGV 121  
 QY 124 KLRVTEGMDDEGHHSESLAYEGRAVDITTSRDRSKYGLMLARLAVENGFDVYYESKAH 183  
 DB 122 KLRVTEGMDDEGHHSESLAYEGRAVDITTSRDRSKYGLMLARLAVENGFDVYYESKAH 181  
 QY 184 IHCVKAKNSVAASGCGFPSSATVHLEHGCTKLVKDLSPEDRYLADADRLYSDEL 243  
 DB 182 IHCVKAKNSVAASGCGFPSSATVHLEHGCTKLVKDLSPEDRYLADADRLYSDEL 241  
 QY 244 FLDMDSRRKLFYVETRPARLLTPAHLFVAPOHNOSEATGTSFG-OALFASANKP 302  
 DB 242 FLDMDSRRKLFYVETRPARLLTPAHLFVAPOHNOSEATGTSFG-OALFASANKP 297  
 QY 303 GQRYVYVAGE--GGQULLPASHVSLREASGAYAPLTAQSTIILNRYLASCYAVIEEHS 360  
 DB 298 GQRYVYVAGE--GGQULLPASHVSLREASGAYAPLTAQSTIILNRYLASCYAVIEEHS 357  
 QY 361 WAHAPAFPRILAGLILALICP-----DGAIPRA-----ATTGTGIMYSLLYRI 405  
 DB 358 WAHAPAFPRILAGLILALILAFARTDGGGGGSIIPAAQSATEARGAEPYAGIHYSOLLYHI 417  
 QY 406 GSWVLDDDALHPLGMVAPAS 425  
 DB 418 GTWLDDSETMHPDGMAYKSS 437

RESULT 4  
 ID SHH\_RAT STANDARD; PRT; 437 AA.  
 AC 063673;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sonic hedgehog protein precursor (SHH).  
 GN SHH OR VHH-1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SURAIN-Sprague-Dawley; TISSUE=Embryonic floor plate;  
 RA MEDLINE=94170375; PubMed=8124714;  
 RA Koellink H., Augsburg A., Heemskerk J., Korzh V., Norlin S.,

RA Ruiz i Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,  
 RA Dodd J.,  
 RA "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog  
 of hedgehog expressed by the notochord.";  
 RL Cell 76:761-775(1994).  
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN  
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION  
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE  
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER  
 CC TARGET, THE G11 ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A  
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED  
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE  
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE  
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH  
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD  
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS  
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC NO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE NODE, NOTOCHORD, FLOOR PLATE,  
 CC AND POSTERIOR LIMB BUD MESENCHYME.  
 CC -1- PFM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L27340; AAA20999.1; -  
 CC HSP: O62226; 1VHH.  
 CC MEROPS: C46.002; -  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hlnt.  
 DR InterPro: IPR003586; Hedgehog\_hlntC.  
 DR InterPro: IPR003587; Hedgehog\_hlntN.  
 DR InterPro: IPR002203; Intein.  
 DR InterPro: IPR001657; SonichH.  
 DR Pfam: PF01079; Hlntc.1.  
 DR Pfam: PF01085; HH\_signal.1.  
 DR PRINTS: PR00632; SONICHOG.  
 DR PRODOM: PD003042; HH\_signal.1.  
 DR SMART: SM00305; Hlntc.1.  
 DR SMART: SM00306; HlntN.1.  
 DR PROSITE: PS50817; INTEIN\_N\_TER.1.  
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KM Signal; Lipoprotein; Palmitate.  
 FT CHAIN 1 24 POTENTIAL.  
 FT CHAIN 25 437 SONIC HEDGEHOG PROTEIN.  
 FT CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.  
 FT CHAIN 199 437 SONIC HEDGEHOG PROTEIN C-PRODUCT.  
 FT SITE 198 199 CLEAVAGE (AUTO-).  
 FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY  
 FT SITE 244 244 SIMILARITY).  
 FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY  
 FT ACT\_SITE 271 271 SIMILARITY).  
 FT BINDING 198 198 ESSENTIAL FOR AUTO-CLEAVAGE (BY  
 FT CHOLESTEROL (BY SIMILARITY).





```

DB 1 MVAATQSLLSLFFICTLTPGACGPGNGICKRRHPKLTPLAYKOFIPNVAEKTIGAS 60
OY 63 GRVEGKITNSSEKFLTPNYNDPIIFKDEENTMGADRLMTOCKDKLNALATSYNMQMG 122
DB 61 GRVEGKITNSSEKFLTPNYNDPIIFKDEENTMGADRLMTOCKDKLNALATSYNMQMG 120
OY 123 VLKRVTEGMEDEGHHSESLHVEGRAVDITTSRDRSKYGMALRLAVEAGFDVVEYSKA 182
DB 121 VLKRVTEGMEDEGHHSESLHVEGRAVDITTSRDRSKYGMALRLAVEAGFDVVEYSKA 180
OY 183 HHCSTKAENSVAAKSGCGFPASATVHLHGGTKLYKDISPDGRLVLAADGRLYSDFL 242
DB 181 HHCSTKAENSVAAKSGCGFPASATVHLHGGTKLYKDISPDGRLVLAADGRLYSDFL 240
OY 243 TFLDRMSRKLFEVYETEPQOPARLLTAAHLLEFVARPHNOSPEATGTSGOALFASNVNP 302
DB 241 MFIIDRDKVKKLFYEYETISQKRLR--LTAHLLLEFVA-----QTKVNGTSEKSVFASNIQD 294
OY 303 GQRVVYLAGGGQOLLPASVSVSLREASGAYAPLTAOGTILINRYLASCYAVIEHSMA 362
DB 295 GDLIYADPKTMTLKAVKEVDL--BEDTGAVAPLTAHGVVVDQVLASCYAVIEHSMA 353
OY 363 HMAFAFLAAGLALALCDGALPTAT-----TTGG 394
DB 354 HMAFAFLAAGLALALCDGALPTAT-----TTGG 394
OY 395 IHWSRLYRIGSWLDDGALHPLGMVAPAS 425
DB 414 IHWSRLYRIGSWLDDGALHPLGMVAPAS 444

```

RESULT 7  
SHH\_BRARE  
ID SHH\_BRARE STANDARD: PRT: 418 AA.  
AC 092008: 013170: 013171;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-OCT-2001 (Rel. 40, Last annotation update)  
DE Sonic hedgehog protein precursor (SHH) (VNH-1).  
GN SHH OR VNH1.  
OS Brachydanio rerio (zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID:7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Embryo;  
RA MEDLINE-94170375; PubMed-8124714;  
RA Roelink H., Augsburg A., Heemskerk J., Korzh V., Norlin S.,  
RA Rold J., Altava A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,  
RA Dodd J.;  
RT "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog  
RT of hedgehog expressed by the notochord";  
RT Cell 76:761-775(1994).  
RN [2]  
RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.  
RX MEDLINE-96014264; PubMed-7583153;  
RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P., Porter J.A.,  
RA Moon R.T., Beachy P.A.;  
RT "Patterning activities of vertebrate hedgehog proteins in the  
RT developing eye and brain";  
RT Curr. Biol. 5:944-955(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-96083328; PubMed-7579523;  
RA Fietz M.J., Concochet J.-P., Barbosa R., Johnson R., Krauss S.,  
RA McMahon A.P., Tablin C., Ingham P.W.;  
RT "The hedgehog gene family in Drosophila and vertebrate development";  
RT development suppl. 43-51(1994).  
RN [4]  
RP SEQUENCE FROM N.A.

```

RX MEDLINE-99225458; PubMed-10207136;
RA Muller F., Chang B., Albert S., Fischer N., Tora L., Strahle U.;
RT "Intrinsic enhancers control expression of zebrafish sonic hedgehog in
RT floor plate and notochord";
RT development 126:2103-2116(1999).
RN [5]
RP SEQUENCE OF 30-92 AND 113-170 FROM N.A.
RC TISSUE-Muscle;
RX MEDLINE-97075114; PubMed-8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
RT closely related to the zebrafish";
RT Proc. Natl. Acad. Sci. U.S.A. 93:13036-13044(1996).
RL
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
CC NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
CC OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
CC FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
CC PTC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
CC TUBE AND BRAIN. ALSO FOUND IN THE NOTOCHORD AND IN DEVELOPING FIN
CC BUD. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN DOMAINS THAT
CC INCLUDE A DISCRETE REGION IN THE FLOOR OF THE Diencephalon.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE INNER CELL LAYER OF
CC THE EMBRYONIC SHIELD DURING GASTRULATION. BY 9.5 HRS OF
CC DEVELOPMENT, EXPRESSED IN A CONTINUOUS BAND THAT EXTENDS FROM THE
CC TAIL TO THE HEAD, THE ANTERIOR BOUNDARY OF EXPRESSION BEING
CC POSITIONED IN THE CENTER OF THE ANIMAL POLE ANTERIOR TO THE
CC PRESUMPTIVE MIDBRAIN.
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L27585; AAA20998.1; -
DR EMBL: U30711; AAC59742.1; -
DR EMBL: Z35669; CAA84738.1; -
DR EMBL: AF124382; AAD47913.1; -
DR EMBL: U51351; AAB38575.1; -
DR EMBL: U51370; AAB38593.1; -
DR HSSP: Q62226; IVNH.
DR MEROPS: C46.002; -
DR ZFIN: ZDB-GENE-980526-166; shh.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001767; Hedgehog_hint.
DR InterPro: IPR003586; Hedgehog_hintc.
DR InterPro: IPR003587; Hedgehog_hintn.
DR InterPro: IPR002203; Intein.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01079; Hint; 1.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PRO0632; SONICHHOG.

```

DR ProDom: PD003042; HH\_signal; 1.  
 DR SMART: SM00305; HintC; 1.  
 DR SMART: SM00306; HintC; 1.  
 DR PROSITE: PS50817; INTERIN\_N\_TER; 1.  
 KM Developmental protein: Autocatalytic cleavage: Hydrolyase: Protease:  
 KW Signal: Lipoprotein: Palmitate.  
 FT CHAIN 1 23  
 FT CHAIN 24 418  
 FT CHAIN 24 197  
 FT CHAIN 198 418  
 FT SITE 197 198  
 FT SITE 243 243  
 FT SITE 267 267  
 FT ACT\_SITE 270 270  
 FT BINDING 197 197  
 FT LIPID 24 24  
 SQ SEQUENCE 418 AA; 46402 MW; CF00A0FFED2F5795 CRC64;  
 Query Match 67.4%; Score 1494; DB 1; Length 418;  
 Best Local Similarity 68.4%; Pred. No. 9.5e-114;  
 Matches 290; Conservative 50; Mismatches 76; Indels 8; Gaps 3;  
 OY 4 MLLTRILLVGFICALLVSSGLTGPGRGIGKRRHPRKLPFLAYKQFIPNVAEKTIGASG 63  
 DB 1 MRLTRVLVSLTLVSLVSGACPGRGRRHPRKLPFLAYKQFIPNVAEKTIGASG 60  
 OY 64 RYEEKITRNSRERFELPNYNPDIIFDEENTGADRLTQCKOKNALAISVANQMPGV 123  
 DB 61 RYEEKITRNSRERFELPNYNPDIIFDEENTGADRLTQCKOKNALAISVANQMPGV 120  
 OY 124 KLRTYEGHDEGHSESLAYEGRAVDITTSRDRSKYGMARLAVEGDFPMVYESKAH 183  
 DB 121 KLRTYEGHDEGHSESLAYEGRAVDITTSRDRSKYGMARLAVEGDFPMVYESKAH 180  
 OY 184 IHCSVKAENSVAAKSGGCFPSATVHLHGGTKLVKDLSPGDVRLADADGRLLYSDEL 243  
 DB 181 IHCSVKAENSVAAKSGGCFPSALVSLDDGOKAVKDLNPDVLAADSNLVFSDPTM 240  
 OY 244 FLDMNDSSRKFLYIETROPARLLTPAHLFLAFOHNSSEAGTSIGQLAFSNAKPG 303  
 DB 241 FTDHDSSTRRFVYIEOEPEYKITTAAHLFLVLDNSTEDLHMT---AAVYSSVAG 296  
 OY 304 QRVVVLSEGGQQLPASVHSVSLREASGAYAPLACSTILINVLASCAVIEHSMVAH 363  
 DB 297 QKVVAVDDSG--LKSIVVQRITTEEDRGSPAPYATAGCTIVYDKILASCAVIEDGLAH 354  
 OY 364 WAFAPFLAOGLLAALCPDGAIPPAAT--TTGIHWYSRLLYRIGSVWLDGDLHPGLMV 421  
 DB 355 LAFAPARLYYVSFLPFQNSSRSNATLQOEGVHWYSRLLYQGTWLDGDMNHPGLGMS 414  
 OY 422 APAS 425  
 DB 415 VNSS 418  
 RESULT 8  
 TWHH\_BRARE  
 ID TWHH\_BRARE STANDARD; PRT; 416 AA.  
 AC Q90419;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Tiggy-winkle hedgehog protein precursor (TWHH).  
 GN TWHH.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]

RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.  
 RC TISSUE-Embryo;  
 RX MEDLINE-96014264; PubMed=7583153;  
 RA Ecker S.C., Ungar A.R., Greenstein P., von Kessler D.P.,  
 RA Potter J.A., Moon R.T., Beachy P.A.;  
 RT "Patterning activities of vertebrate hedgehog proteins in the  
 RT developing eye and brain".  
 RL Curr. Biol. 5:944-955(1995).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT. INVOLVED IN DORSO-VENTRAL  
 CC PATTERNING OF THE BRAIN AND IN EARLY PATTERNING OF THE DEVELOPING  
 CC EYES.  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL  
 CC TUBE AND BRAIN. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN  
 CC DOMAINS THAT INCLUDE A DISCRETE REGION IN THE FLOOR OF THE  
 CC DIENCEPHALON, NOT DETECTED IN THE NOTOCHORD OR DEVELOPING FIN BUD.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR HSPB; U30710; AAC59741.1; -.  
 DR MEROPS; C46.001; -.  
 DR ZFIN; ZDB-GENE-980526-41; twhh.  
 DR InterPro; IPR000320; HH\_signal.  
 DR InterPro; IPR001767; Hedgehog\_hntc.  
 DR InterPro; IPR003586; Hedgehog\_hntc.  
 DR InterPro; IPR003587; Hedgehog\_hntc.  
 DR InterPro; IPR001657; SonICHH.  
 DR Pfam; PF01079; HintC; 1.  
 DR Pfam; PF01085; HH\_signal; 1.  
 DR PRINTS; PRO0632; SONICHOG.  
 DR PRODOM; PD003042; HH\_signal; 1.  
 DR SMART; SM00305; HintC; 1.  
 DR SMART; SM00306; HintN; 1.  
 KW Developmental protein: Autocatalytic cleavage: Hydrolyase: Protease;  
 KW Signal: Lipoprotein: Palmitate.  
 FT SIGNAL 1 26  
 FT CHAIN 27 416  
 FT CHAIN 201 416  
 FT CHAIN 201 416  
 FT SITE 200 201  
 FT SITE 270 270  
 FT ACT\_SITE 273 273  
 FT BINDING 200 200  
 FT LIPID 27 27  
 SQ SEQUENCE 416 AA; 46576 MW; 61EC2218309CFE59 CRC64;  
 Query Match 65.7%; Score 1458; DB 1; Length 416;  
 Best Local Similarity 68.8%; Pred. No. 7.9e-111;  
 Matches 287; Conservative 36; Mismatches 86; Indels 8; Gaps 4;

QY 5 LILTRLLVGFICALLVSSGLTGPGRGIGKRRHPRKPLPLAYKOFIPNVAEKTLAGSGR 64  
 DB 5 LHKOPALLCFILSLTPCLACGPGRGYGRHPRKPLPLAYKOFIPNVAEKTLAGSGR 64  
 QY 65 YEKTRNSRERKELTPNPNPDIIKDEENTGADRLMORCKKLNALAI SVMNOMPVKLR 124  
 DB 65 YEKTRNSRERKELTPNPNPDIIKDEENTGADRLMORCKKLNALAI SVMNOMPVKLR 124  
 QY 125 LRTYEGMEDGHSESLHYEGRAVDITTSDRSRKYGMRLARLAVAGEFDMVYESKANH 184  
 DB 125 LRTYEGMEDGHSESLHYEGRAVDITTSDRSRKYGMRLARLAVAGEFDMVYESKANH 184  
 QY 185 HCSVKAENSVAAKSGCGPFGSATVHLHSGTKLVKDLSPEDRYLAADADGRLLYSDFLAF 244  
 DB 185 HCSVKAENSVAAKSGCGPFGSATVHLHSGTKLVKDLSPEDRYLAADADGRLLYSDFLAF 244  
 QY 245 LDMDSRKLFFVYIETROPARLLTAHLFVAPOHNOSEATGSTSGOALFASNVKPGQ 304  
 DB 245 LDMDSRKLFFVYIETROPARLLTAHLFVAPOHNOSEATGSTSGOALFASNVKPGQ 304  
 QY 305 RYVVLGEGGQQLLPASVSHVSLREASGAVAPLTAOCTILINRYLASCVAVIEHSHMAHW 364  
 DB 299 TYLVWMDTCSLSKSVYKRI-YTEHEGSPAPYANGTIIVDQVLASCVAIENHMAHW 357  
 QY 365 AFAPRLAOGGLAALCPDGAIPPAATTTGIIHWYSLRLYRIGSWLDGALHPGAV 421  
 DB 358 AFAPRLAOGGLAALCPDGAIPPAATTTGIIHWYSLRLYRIGSWLDGALHPGAV 421  
 RESULT 9  
 ID IHH\_CHICK STANDARD; PRT; 408 AA.  
 AC 098938;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Indian hedgehog protein precursor (IHH).  
 GN IHH.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP MEDLINE=96325423; PubMed=8662546;  
 RA Vortkamp A., Lee K., Lanske B., Segre G.V., Kronenberg H.M.,  
 RA Tabin C.J.;  
 RT Regulation of rate of cartilage differentiation by Indian hedgehog  
 RT and PTH-related protein.  
 RL Science 273:613-622(1996).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)  
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO  
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN  
 CC ENDOCHONDRAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH  
 CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION  
 CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP).  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN DEVELOPING MIDGEUT, LUNG AND  
 CC CARTILAGE OF DEVELOPING LONG BONES IN THE LIMB.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOBIOLOGOUS ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 CC EMBL: U58511; AAC60010.1; -  
 CC HSP: Q62226; 1VNH.  
 DR MEROPS: C46.003; -  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hlnc.  
 DR InterPro: IPR003586; Hedgehog\_hlnc.  
 DR InterPro: IPR003587; Hedgehog\_hlnc.  
 DR InterPro: IPR01657; SonicHH.  
 DR Pfam: PF01079; Hlnc.1.  
 DR Pfam: PF01085; HH\_signal.1.  
 DR PRINTS: PR00632; SONICHHOG.  
 DR ProDom: PD003042; HH\_signal.1.  
 DR SMART: SM00305; Hlnc.1.  
 DR SMART: SM00306; Hlnc.1.  
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KW Signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 23  
 FT CHAIN 24 408  
 FT CHAIN 24 198  
 FT CHAIN 199 408  
 FT SITE 198 408  
 FT SITE 244 244  
 FT SITE 244 244  
 FT SITE 268 268  
 FT ACT\_SITE 271 271  
 FT BINDING 198 198  
 FT LIPID 24 24  
 SQ SEQUENCE 408 AA; 44829 MW; BA397AE2A0357A24 CRC64;  
 Query Match 58.4%; Score: 1295.5; DB 1; Length 408;  
 Best local Similarity 64.3%; Pred. No. 1.2e-97;  
 Matches 266; Conservative 39; Mismatches 90; Indels 19; Gaps 7;  
 QY 9 RILLVGFICALLVSSGLTGPGRGIGKRRHPRKPLPLAYKOFIPNVAEKTLAGSGR 66  
 DB 5 RILLVGFICALLVSSGLTGPGRGIGKRRHPRKPLPLAYKOFIPNVAEKTLAGSGR 66  
 QY 67 GKTTRNSRERKELTPNPNPDIIKDEENTGADRLMORCKKLNALAI SVMNOMPVKLR 126  
 DB 65 GKTTRNSRERKELTPNPNPDIIKDEENTGADRLMORCKKLNALAI SVMNOMPVKLR 126  
 QY 127 VTEGMDGHSESLHYEGRAVDITTSDRSRKYGMRLARLAVAGEFDMVYESKANHIC 186  
 DB 125 VTEGMDGHSESLHYEGRAVDITTSDRSRKYGMRLARLAVAGEFDMVYESKANHIC 186  
 QY 187 SYKAEKNSVAAKSGCGPFGSATVHLHSGTKLVKDLSPEDRYLAADADGRLLYSDFLAF 244  
 DB 185 SYKAEKNSVAAKSGCGPFGSATVHLHSGTKLVKDLSPEDRYLAADADGRLLYSDFLAF 244  
 QY 247 RMDSRKLFFVYIETROPARLLTAHLFVAPOHNOSEATGSTSGOALFASNVKPGQ 306  
 DB 245 KEPRALTAHFVETROPARLLTAHLFVAPOHNOSEATGSTSGOALFASNVKPGQ 300  
 QY 307 YVVLGEGGQQLLPASVSHVSLREASGAVAPLTAOCTILINRYLASCVAVIEHSHMAHW 365  
 DB 301 LVAVGSGGLD--PAEYVGRGRDY--GAYAPLTHGTLVVDVAVSCFALVEEQLOAOWA 357  
 QY 366 AFAPRLAOGGLAALCPDGAIPPAATTTGIIHWYSLRLYRIGSWLDGALHPGAV 419  
 DB 358 FWPRLYHSLDGGPGVQ-----DGVHWYSGLLYRGLRPLPDSFHPG 402





RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,  
 RA McMahon J.A., McMahon A.P.;  
 RT "Sonic hedgehog, a member of a family of putative signaling  
 RT molecules, is implicated in the regulation of CNS polarity.";  
 RL Cell 75:1417-1430(1993).  
 RN [3]  
 RP REVISIONS.  
 RC STRAIN=C57BL/6J;  
 RA St Jacques B.;  
 RL Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 124-172 FROM N.A.  
 RX MEDLINE=95236997; Pubmed=7720571;  
 RA Chang D.T., Lopez A., von Kessler D.P., Chang C., Simandl B.K.,  
 RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;  
 RT "Products, genetic linkage and limb patterning activity of a murine  
 RT hedgehog gene";  
 RL Development 120:3339-3353(1994).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)  
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO  
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN  
 CC ENDOCHONDRAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH  
 CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION  
 CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP).  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, DETECTED IN THE DEVELOPING GUT,  
 CC AND UROGENITAL SINUS. IN THE ADULT KIDNEY, FOUND IN PROXIMAL  
 CC CONVOLUTED AND PROXIMAL STRAIGHT TUBULE.  
 CC -1- DEVELOPMENTAL STAGE: DETECTED AT 10 DAYS POST COITUM (DPC) IN  
 CC IN THE DEVELOPING UROGENITAL SINUS. EXPRESSION INCREASES WITH  
 CC GESTATIONAL AGE IN KIDNEY AND DUODENUM, BECOMING MAXIMAL IN  
 CC ADULTHOOD.  
 CC -1- THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL: U85610; AAB49692.1; ALT\_INIT.  
 CC EMBL: X76291; CA53923.1; --  
 CC HSSP: 062226; 1VHH.  
 CC DR MEROPS: C46.003; --  
 CC DR MGD: MGI:96533; Inh.  
 CC DR InterPro: IPR000320; HH\_signal.  
 CC DR InterPro: IPR001767; Hedgehog\_hint.  
 CC DR InterPro: IPR003586; Hedgehog\_hintc.  
 CC DR InterPro: IPR003587; Hedgehog\_hintn.  
 CC DR InterPro: IPR002203; Intein.  
 CC DR Pfam: PF01079; Hinc; 1.  
 CC DR Pfam: PF01085; HH\_signal; 1.  
 CC DR ProDom: PD003042; HH\_signal; 1.  
 CC DR SMART: SM00305; Hinc; 1.  
 CC DR SMART: SM00306; Hinc; 1.  
 CC DR PROSITE: PSS0817; INTEIN\_N\_TER; 1.

KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KM Signal: Lipoprotein; Palmitate.  
 FT SIGNAL 1 27  
 FT CHAIN 28 411  
 FT CHAIN 28 202  
 FT CHAIN 203 411  
 FT SITE 202 203  
 FT SITE 202 203  
 FT SITE 202 248  
 FT SITE 272 272  
 FT ACT\_SITE 275 275  
 FT BINDING 202 202  
 FT LIPID 28 28  
 FT CONFLICT 383 383  
 FT SEQUENCE 411 AA; 45485 MW; 08BEFADB507C00DB CRC64;  
 SO SEQUENCE  
 Query Match 57.68; Score 1277.5; DB 1; Length 411;  
 Best Local Similarity 62.48; Pred. No. 3,4e-96;  
 Matches 257; Conservative 42; Mismatches 88; Indels 25; Gaps 6;  
 Oy 19 LVSSGLTCCGRCIG-KRRHPRKLTPLAYKQIPNVAERTKLSGVEGKITNSERK 77  
 Db 20 LTVPAKCGCGRVGSSRRPRKLVPLAKQSPNPEKTLGASGREKIASSERFK 79  
 Oy 78 ELTPNYPNDIFKDEMGADRLTQRCOKDLNLAISVNMQGVKLVTEGDEGSH 137  
 Db 80 ELTPNYPNDIFKDEMGADRLTQRCOKDLNLAISVNMQGVKLVTEGDEGSH 139  
 Oy 138 SEESLHTEGRAVDITTSDDRSKTYGMLARLAVEGFMVYTESFAHICSVKAENVA 197  
 Db 140 SEESLHTEGRAVDITTSDDRSKTYGMLARLAVEGFMVYTESFAHICSVKAENVA 199  
 Oy 198 SGGEFPGSAFVHLHGHTKTKYKDLSPDRVLAADRLYSDTLTDRDSSKLFY 257  
 Db 200 TGGEFPGAGVRLNGERVALAVKPGDRVLADEGDPPTFSVLTFLDREPNLRV 259  
 Oy 258 IETQPRARLLTAAHLFLVAPQHNQSEATSTGQALFASNVPGORVYVLEGGOQL 317  
 Db 260 IETQPRARLLTAAHLFLVAPQHNQSEATSTGQALFASNVPGORVYVLEGGOQL 313  
 Oy 318 PASVHVSFLREASGAVAPLTAQSTILINRYLASCIVIEHSAHNAEPFRLAQGLA 377  
 Db 314 PARVAASV-THVALGSYAPLFRHGTLVVEDVVASCFPAVADHHAQLAFPLRL- 366  
 Oy 378 ALCDGAIPTAA-----PTTGIHVSRLYRIGSVWLDGALHPLGVAPAS 425  
 Db 367 -----FPLAMGSMTPBSGVMTYPMWLRGLLLEESTFPLGSGAGS 411  
 RESULT 12  
 ID IHH\_HUMAN STANDARD; PRT; 411 AA.  
 AC Q14623; O43322;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Indian hedgehog protein precursor (IHH) (HHG-2).  
 GN IHH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tate G., Kishimoto K., Mitsuya T.;  
 RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in  
 RT human cancer cell lines and embryonic organs";  
 RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).  
 RN [2]  
 RP SEQUENCE OF 100-411 FROM N.A.  
 RC TISSUE=Fetal lung;



RX MEDLINE-96070431; PubMed-7590746;  
 RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levy T.,  
 RA Gaster J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seidman C.E.,  
 RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.;  
 RT "Cloning, expression, and chromosomal location of SHN and HH: two  
 RT human homologues of the Drosophila segment polarity gene hedgehog.";  
 RN Genomics 28:44-51(1995).  
 RP SEQUENCE OF 124-172 FROM N.A.  
 RX MEDLINE-95236997; PubMed-7720571;  
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,  
 RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;  
 RT "Products, genetic linkage and limb patterning activity of a murine  
 RT hedgehog gene.";  
 RL Development 120:3339-3353(1994).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)  
 CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO  
 CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN  
 CC ENDOCHONDRAL OSSIFICATION. MAY REGULATE THE BALANCE BETWEEN GROWTH  
 CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION  
 CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP) (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC LUNG, AND IN ADULT  
 CC KIDNEY AND LIVER.  
 CC -1- PM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING. WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; AB018076; BAA33523.2;  
 DR EMBL; AB010092; BAA33523.2; JOINED.  
 DR EMBL; AB018075; BAA33523.2; JOINED.  
 DR EMBL; L38517; AAA62178.1;  
 DR HSSP; Q62226; 1VHH.  
 DR MEROPS; C46.003;  
 DR GeneW; HGNC:5956; HH.  
 DR MIM; 600726;  
 DR InterPro; IPR000320; HH\_signal.  
 DR InterPro; IPR001767; Hedgehog\_hnt.  
 DR InterPro; IPR003586; Hedgehog\_hntc.  
 DR InterPro; IPR003587; Hedgehog\_hntn.  
 DR InterPro; IPR002203; Inteln.  
 DR Pfam; PF01079; Hnt; 1.  
 DR Pfam; PF01085; HH\_signal; 1.  
 DR SMART; SM00305; Hntc; 1.  
 DR SMART; SM00306; Hntn; 1.  
 DR PROSITE; PS0817; INTERIN\_N\_TER; 1.  
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KW Signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 27  
 FT CHAIN 28 411  
 FT CHAIN 28 202  
 FT CHAIN 203 411  
 FT SITE 202 203  
 FT SITE 248 248  
 FT SITE 248 248  
 INVOLVED IN CHOLESTEROL TRANSFER (BY

FT SITE 272 272  
 FT ACT\_SITE 275 275  
 FT BINDING 202 202  
 FT LIPID 28 28  
 FT CONFLICT 100 100  
 FT CONFLICT 246 246  
 FT CONFLICT 309 309  
 SQ SEQUENCE 411 AA; 4518 MW; 14C90436C344DEA CRC64;  
 Query Match 57.1%; Score 1265.5; DB 1; Length 411;  
 Best Local Similarity 62.7%; Pred. No. 3.2e-95;  
 Matches 256; Conservative 38; Mismatches 97; Indels 17; Gaps 5;  
 QY 19 LTVSSGLTCGPRGIG-KRRHKRLPLAYKQPTRNAVKETIGASGRREGKTRNSERK 77  
 DB 20 LVVPAMGCGPGRVYSSRRPRKLVPLAYKQFSNVPEKTIGASGRREGKTRNSERK 79  
 QY 78 ELTPVNPDIIRKDEBNTGADRLMTORCKDLNALAISVMNMPGVKLVTEGMEDGH 137  
 DB 80 ELTPVNPDIIRKDEBNTGADRLMTORCKDLNALAISVMNMPGVKLVTEGMEDGH 139  
 QY 138 SESEHYEGRAVDITTSDRSKYGLARLAVEAGFDWYYSKAKIHCSVAENSVAK 197  
 DB 140 SESEHYEGRAVDITTSDRSKYGLARLAVEAGFDWYYSKAKIHCSVAENSVAK 199  
 QY 198 SCGCPGSGATVLENGRTKLVNDLSPGRVNLADADGRLYSDFLTLDMDSSKRLPV 257  
 DB 200 TGCCPAGAGVLESGARVALSARVGRDVLAMGDSPTFSDVLTLDRPHRRRAFOY 259  
 QY 258 ITTPPRARLTAHLLEFAPHOHNSSEATGSGOALFASNVKRGORVYVGEQQQL 317  
 DB 260 ITTPPRARLTAHLLEFAPHOHNSSEATGSGOALFASNVKRGORVYVGEQQQL 319  
 QY 318 PASVHSVLSREBSAGAVAPLTAQGTILNRLVLAQVAVTEESHMAWAFAPRLAOGILA 377  
 DB 314 PARVAAYS-THVALGAVAPLTKHGTLLVEDVAVSCFAAVADHNLQALWPLRFLSHLW 372  
 QY 378 ALCPDGAIPATATTTGTHWYSRLXLRIGSWLDGALPCLGVAPAS 425  
 DB 373 G-----SWTPEGVMYPPOLLRLGLRLLESGFHPIDMSGNS 411  
 RESULT 13  
 IHH\_BRARE  
 AC 098862; STANDARD; PRT; 412 AA.  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Indian hedgehog protein precursor (IHH) (Echidna hedgehog protein).  
 GN IHH OR EHH.  
 OS Brachydanio rerio (zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_Taxid=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-96310864; PubMed-8684485;  
 RA Currie P.D., Ingham P.W.;  
 RT "Induction of a specific muscle cell type by a hedgehog-like protein  
 RT in zebrafish.";  
 RL Nature 382:452-455(1996).  
 RN [2]  
 RP SEQUENCE OF 113-170 FROM N.A.  
 RC TISSUE=Muscle;  
 RX MEDLINE-97075114; PubMed-8917540;  
 RA Zardoya R., Abouheif E., Meyer A.;  
 RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
 RT closely related to the zebrafish.";

```

RL  Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC  -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC  PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
CC  NOTOCHORD THAT INDUCES SOMITE PATTERNING AND MUSCLE PIONEER
CC  DIFFERENTIATION.
CC  -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC  CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC  CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC  TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC  -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN THE NOTOCHORD.
CC  -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT THE MID-GASTRULA STAGE.
CC  DISAPPEARS AT THE END OF THE SOMITOGENESIS.
CC  -1- PFM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC  AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC  THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC  A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC  TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC  TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC  OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC  ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC  C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC  -----
CC  THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC  BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
CC  THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC  USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC  MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC  ENTITIES REQUIRES A LICENSE AGREEMENT (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; Y08426; CAA69702.1; -.
DR  HSSP; O62226; 1VHN.
DR  MEROPS; C46.0PM; -.
DR  ZFIN; ZDB-GENE-980526-135; enh.
DR  InterPro; IPR000320; HH_signal.
DR  InterPro; IPR001767; Hedgehog_hntc.
DR  InterPro; IPR003586; Hedgehog_hntc.
DR  InterPro; IPR003587; Hedgehog_hntc.
DR  InterPro; IPR002203; Intein.
DR  InterPro; IPR001657; SonicHH.
DR  Pfam; PF01079; Hntc; 1.
DR  Pfam; PF01085; HH_signal; 1.
DR  PRINTS; PR006332; SONICHHOG.
DR  ProDom; PD003042; HH_signal; 1.
DR  SMART; SM00305; Hntc; 1.
DR  SMART; SM00306; Hntn; 1.
DR  PROSITE; PSS0817; INTEIN_N_TERM; 1.
KM  Developmental protein. Autocatalytic cleavage; Hydrolase; Protease;
KM  Signal; Lipoprotein; Palmitate.
FT  SIGNAL 1 23
FT  CHAIN 24 412
FT  CHAIN 24 197
FT  CHAIN 198 412
FT  SITE 197 198
FT  SITE 246 246
FT  SITE 269 269
FT  ACT_SITE 272 272
FT  BINDING 197 197
FT  LIPID 24 24
FT  CONFLECT 121 121
SO  SEQUENCE 412 AA; 45746 MW; 21D43F052A278CA1 CRC64;

Query Match 55.68; Score 1232.5; DB 1; Length 412;
Best Local Similarity 60.98; Pred. No. 1.5e-92;
Matches 254; Conservative 51; Mismatches 91; Indels 21; Gaps 10;

OY 11 LVAFGICALVSSGSGGPGGIGRRHPRKPLPLAYKOFIPNVAETIASGVEEKIT 70
DB 9 LITGFIILAFSPAYD-GGPGGIGRRKRTPKRLPLAYKOPSPNVAETILASGVEEKIT 67

```

```

OY 71 RNSRFKELPNYNDIIFDEENTGADRLMTQRCXKDLNALAISVNMOPGVRLRYTEG 130
DB 68 PSSRFKELPNYNDIIFDEENTGADRLMTQRCXKDLNALAISVNMOPGVRLRYTEG 127
OY 131 WDEGHHSESLHREGAVDITTSDRORSKYGMALRLAVGEPWVYTESKAHHCYKA 190
DB 128 WDEGHHSESLHREGAVDITTSDRORSKYGMALRLAVGEPWVYTESKAHHCYKA 187
OY 191 ENSVAKSGGCPSPAVHLEHGSTRKYNDLSPGDRVLA---ADADGRLLYSDPLFLDR 247
DB 188 EHSVAATGGCFERBALVTKDGSROIQDAGDLVASGSGTGDILYEVLTFLDR 247
OY 248 MDSRKLVEYLETROPARLLTAHLIF--VAPDHNQSEATGTSQALFASVYKPGOR 305
DB 248 RPIQKHFVYIRT-EDGASVSLTAHLIFKRVGNSKRGEPKG-AVRTIFASDAQGQ- 304
OY 306 VYVLEGGQQLPASVSVSLREASGAYAPLTAOGTILNRLVASCYAVIEEHSNAHMA 365
DB 305 CLTGKLRKR--SQTHVGRED-QGLYPLTAHGTIVVNDVLTSCYAAVNRQRLAHMA 361
OY 366 FAPRLAGLIALCPGAIPTATTTGTGHWYSRLYRIGSWVLDGDLHPGMA 422
DB 362 FAPRL--LYSWGPDQVL-----KNGLHWISQVLIGKLLDSELPHPALAE 409

RESULT 14
ID DHH_MOUSE STANDARD; PRT; 396 AA.
AC 061488;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Desert hedgehog protein precursor (DHH) (HHG-3).
GN DHH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_TaxID=10090;
RN 1[1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=94094334; Pubmed=7916661;
RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,
RT McMahon J.A., McMahon A.P.;
RT "Sonic hedgehog, a member of a family of putative signaling
RT molecules, is implicated in the regulation of CNS polarity.";
RL Cell 75:1417-1430(1993).
RN 12[2]
RP SEQUENCE OF 120-168 FROM N.A.
RX MEDLINE=95236997; Pubmed=7720571;
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,
RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene.";
RL Development 120:3339-3353(1994).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A
CC SPERMATOCTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES
CC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTES. NOT EXPRESSED IN
CC LIMB BUDS.
CC -1- PFM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE

```

CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 DR EMBL: X76292; CA53924.1; -  
 DR HSSP: Q62226; LVH.  
 DR MEROPS: C46.004; -  
 DR MGD: MGI:94891; Dh.  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hnt.  
 DR InterPro: IPR003586; Hedgehog\_hntc.  
 DR InterPro: IPR003587; Hedgehog\_hntn.  
 DR InterPro: IPR002203; Intein.  
 DR Pfam: PF01079; Hnt; 1.  
 DR Pfam: PF01085; HH\_signal; 1.  
 DR PRINTS: PR00632; SONICHOG.  
 DR PRODOM: PD003042; HH\_signal; 1.  
 DR SMART: SM00305; Hntc; 1.  
 DR SMART: SM00306; Hntn; 1.  
 DR PROSITE: PSS0817; INTEIN\_N\_TER; 1.  
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KW Signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 22  
 FT CHAIN 23 396  
 FT CHAIN 23 198  
 FT CHAIN 199 396  
 FT SITE 198 199  
 FT SITE 244 244  
 FT SITE 268 268  
 FT ACT\_SITE 271 271  
 FT BINDING 198 198  
 FT LIPID 23 23  
 FT SEQUENCE 396 AA; 43542 MW; AFEB051BE950FD8 CRC64;  
 SO  
 Query Match 54.18; Score 1200; DB 1; Length 396;  
 Best Local Similarity 60.78; Pred. No. 6.3e-90;  
 Matches 241; Conservative 56; Mismatches 84; Indels 16; Gaps 9;  
 QY 12 LVGFTIC-ALVSSGLTGCGRNG-IGKRRHPRK-LTPLAYKQFIPNVAEKTLAGSGRYEGK 68  
 Db 7 LPLPCCLALLALASOSGPGRGPRVGRYVKOLVPLIKYFVSPMERITGASGPACRG 66  
 QY 69 ITRNSERKELTPYNDPIIKDEBNTGADRLMQRCKDKLNALAISSVNMOPGVKRLVT 128  
 Db 67 VTRNSERRDLVPPNYNDPIIKDEBNTGADRLMQRCKDKLNALAISSVNMOPGVKRLVT 126  
 QY 129 EGMDEDGHSSESLHYEGRAVDITTSDDRSKYGMLARLAEGAPDWMYSEKAHIHOSV 188  
 Db 127 EGMDEDGHSSESLHYEGRAVDITTSDDRSKYGMLARLAEGAPDWMYSEKAHIHOSV 186  
 QY 189 KAENSVAKSGCGPPGASVTVHLEHGCTKLVNDLSPGDVLAADADGRLLVSDPLTFIDRM 248  
 Db 187 KADNSLAVRAGCGPPGASVTVHLEHGCTKLVNDLSPGDVLAADADGRLLVSDPLTFIDRM 246  
 QY 249 DSSKRLTYVITRPPARLLTAAHLFVAPQHNOSEATGTSQALFASVAVKPGQRYTV 308  
 Db 247 LQRAESFAVETERRPPRLTLTPMHLVFAA--RGAPAPGCDFA--PVFARRLRAGDS--V 300  
 QY 309 LGEAGDGLLPASVSVLSREASAGVAPLTAGQTLIRVASCYAVVEEESHMHWAP 368  
 Db 301 LAPGGLALQPARVARVA-REZAVGVFALTAHGLLVNDVLASCYAVVEEESHMHWAP 359

QY 369 FRLAAGLALCPDGAIPTAATTGTHMYSRLRYR 405  
 Db 360 LRLHA-LGALLPGAV-----QPTGMHWYSRLRYRL 390  
 RESULT 15  
 DHH\_HUMAN  
 ID DHH\_HUMAN STANDARD; PRT; 396 AA.  
 AC 043323; 015794;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Desert hedgehog protein precursor (DHH) (HNG-3).  
 GN DHH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID:9606;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RA Tate G., Kishimoto K., Mitsuya T.;  
 RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in  
 RT human cancer cell lines and embryonic organs";  
 RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).  
 RN [2]  
 RP SEQUENCE OF 85-178 FROM N.A.  
 RC TISSUE=Kidney;  
 RA Drummond I.A.;  
 RT "Human desert hedgehog";  
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP DISEASE.  
 RX PubMed:11017805;  
 RA Umehara F., Tate G., Itoh K., Yamaguchi N., Douchi T., Mitsuya T.,  
 RA Osame M.;  
 RT "A novel mutation of desert hedgehog in a patient with 46,xy partial  
 RT gonadal dysgenesis accompanied by multifocal neuropathy";  
 RL Am. J. Hum. Genet. 67:1302-1305(2000).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL. ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A  
 CC SPERMATOCTYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES  
 CC DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- DISEASE: Defects in DHH may be the cause of partial gonadal  
 CC dysgenesis (PGD) with multifocal neuropathy. PGD is  
 CC characterized by the presence of a testis on one side and a streak  
 CC or an absent gonad at the other, persistence of Mullerian duct  
 CC structures, and a variable degree of genital ambiguity.  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 DR EMBL: AB010994; BAA24866.1; -  
 DR EMBL: AB010581; BAA24866.1; JOINED.

EMBL: AB010993; BAA24866.1; JOINED.  
DR EMBL: U59748; AAB03398.1; -.  
DR HSSP: Q62226; 1VHH.  
DR MEROPS: C46.004; -.  
DR Genew: HGNC:2865; DHH.  
DR MIM: 605423; -.  
DR MIM: 607080; -.  
DR InterPro: IPR000320; HH\_signal.  
DR InterPro: IPR001767; Hedgehog\_hntc.  
DR InterPro: IPR003586; Hedgehog\_hntc.  
DR InterPro: IPR003587; Hedgehog\_hntc.  
DR InterPro: IPR001657; SonicHH.  
DR Pfam: PF01079; Hntc; 1.  
DR Pfam: PF01085; HH\_signal; 1.  
DR PRINTS: PR00632; SONICHHOG.  
DR ProDom: PD003042; HH\_signal; 1.  
DR SMART: SM00305; Hntc; 1.  
DR SMART: SM00306; Hntc; 1.  
DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
KW Signal; Lipoprotein; Palmitate.  
KW Signal; Lipoprotein; Palmitate.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 396 DESERT HEDGEHOG PROTEIN.  
FT CHAIN 23 198 DESERT HEDGEHOG PROTEIN N-PRODUCT.  
FT CHAIN 199 396 DESERT HEDGEHOG PROTEIN C-PRODUCT.  
FT SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).  
FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY  
SIMILARITY).  
FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY  
SIMILARITY).  
FT ACT\_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY  
SIMILARITY).  
FT BINDING 198 198 CHOLESTEROL (BY SIMILARITY).  
FT LIPID 23 23 PALMITATE (BY SIMILARITY).  
FT CONFLICT 177 177 E -> G (IN REF. 2).  
SQ SEQUENCE 396 AA: 43577 MW: FCE4FB21972C3AD5 CRC64:

Query Match 54.0%; Score 1198; DB 1; Length 396;  
Best Local Similarity 59.4%; Pred. No. 9, 2e-90;

Matches 244; Conservative 56; Mismatches 91; Indels 20; Gaps 10;

QY 4 MLTFRILLVFGIC--ALVSSGLTGGPGRG-IGRRHPK--LTPLAYKOFIPNVAKTL 59  
DB 1 MALLTFLLP---LCCIALALAPRQSCGPRGPRGRRYARKOLVPLLYKQFVGPVERKL 57  
QY 60 GASGRYEGKITNSERKEKELTPYNPDIIFKDEENTGADRLMTQRCCKLNLAIISVNO 119  
DB 58 GASGPAEGHVARGSERFDLPYNPDIIFKDEENSGADRLMTRECKERYNLAIIVMNM 117  
QY 120 WPGVKLRVTEGDEGHSEESIHYGRAVDITTSRDRSKYGMALARLAVEAGFDWVYE 179  
DB 118 WPGVRLRVTEGDEGHSEESIHYGRAVDITTSRDRSKYGMALARLAVEAGFDWVYE 177  
QY 180 SKAHISCVKAENSVAKSGGCFPGSATVHLEHGTGKLVKDSPGDRVLAADADGRLYS 239  
DB 178 SRNHVIVSVKADNSLAVRAGCGCPGNATVRLNSGERKGLRELHGDVWVLAADSGVVPPT 237  
QY 240 DELTFLDRMDSRKLEFYIETROPARALLTLTAHLLEFVAPHNQSEATSGTSGALFASN 299  
DB 238 PVLLFDRDLQRRASFVAVETEMPRLKLLTPMHLVFAA--RGPAPAPGDFA--PVFARR 293  
QY 300 VKPGQRYVYVLGGGQQLLPASVHVSLEBASGAYAPLTAQCTILINRYLASCYAVIEEH 359  
DB 294 LKAGDS--VLAPGDALRPAPARVA--REAVGVFAPLTHGTLVNDVASCYAVLESH 350  
QY 360 SWAHNAFAPFRLAOGLLALCPDGAIPATAITTTGTHWTSRLLYRIGSWVL 410  
DB 351 QNAHRAFAFLRLIHA-LGALLPGAV-----OPTGHHMTSRLLYRAEELL 395

Search completed: February 20, 2003, 10:12:37  
Job time : 14.0329 secs



Qy	1	MVEMLLTRLLVAFETALVSSLTGCGPGGICKRKKPKPLAYKOTIPVAAETLG	60
Dy	1	MGEMLLRDLRGCFITSLVPLVGLGCGPGRG16QRRRPOKPLAYKOTIPVASEKTG	60
Qy	61	ASGRECKITRNSRFRKELPKNYNDIIFDEENTGADRLMTOCRCKDKNALAISVNMW	120
Dy	61	ASGREYVITRNSRFRQGLPKNYNDIIFKYEENTGADRLMTOCRCKDKLMSLAISVNMW	120
Qy	121	PGVKRLYTEGDEGCHSEESLHIEGNAVDTITSDRDRSKYGMALRLAVAGEDWYTES	180
Dy	121	PGVKRLYTEGDEGCHHSEESLHIEGNAVDTITSDRDRSKYGMALRLAVAGEDWYTES	180
Qy	181	KAHHCISVKAENSVAAKSGGCPGSAIVYHLEHGCTKLVKDLSPGDRLVADAGORLLYS	240
Dy	181	KAHHCISVKAENSVAAKSGGCPGSAIVYHLEHGCTKLVKDLSPGDRLVADAGORLLYS	240
Qy	241	FLVFLDRMDSRKLFEYIETROPARLLTAHLFLFAPOHNSSEATGSTG-----Q	293
Dy	241	FLVFLDEEETRRKFEYIETSLPERLRLTAAHLFLFAQEHGPNASNSFRSKGHRFR	300
Qy	294	ALFASNYKPGCRVYVLEGCGQLLPASVHSLSLREESGNAVPLTACGTLILNRVLASCY	353
Dy	301	SMFASVAPRGHRLVLEDEREGLEAVYDRLYL-EEATGAAPATAGVAVIDRLVASCY	359
Qy	354	AVIEESHVAHNAFAPFRLAQLALACP---DGAIPTAATTTTGHHYSRLYRIGSMVL	410
Dy	360	AVIEESHVAHNAFAPRLKRGFILLFFSPRODSHSPAPSGABEVHYSRLYRIGTWVL	419
Qy	411	DGDALHPLGVNAPAS 425	
Dy	420	QADTIHPLGMAAKSS 434	
RESULT 2			
Q9W709		PRELIMINARY;	PRT; 414 AA.
AC	Q9W709;		
DT	01-NOV-1999 (TREMblrel. 12, Created)		
DT	01-NOV-1999 (TREMblrel. 12, Last sequence update)		
DT	01-MAR-2002 (TREMblrel. 20, Last annotation update)		
DE	Sonic hedgehog.		
GN	SHH.		
OS	Paralichthys olivaceus (Flounder).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;		
OC	Pleuronectoidae; Paralichthyidae; Paralichthys.		
OX	NCBI_TaxID=8255;		
RP	[1]		
RA	SEQUENCE FROM N.A.		
RX	MEDLINE=99238226; PubMed=10223710;		
RT	Suzuki T., Ichiro O., Kurokawa T.;		
RT	"Retinoic acid given at late embryonic stage depresses sonic hedgehog		
RT	and Hoxd-4 expression in the pharyngeal area and induces skeletal		
RT	malformation in flounder (Paralichthys olivaceus) embryos."		
RL	Dev. Growth Differ. 41:141-152(1999).		
DR	EMBL: AB029748; BAA82360.1; -		
DR	HSSP: O62226; IVHH.		
DR	MEROPS: C46.001;		
DR	InterPro: IPR001767; Hedgehog_hlnt.		
DR	InterPro: IPR003586; Hedgehog_hlntc.		
DR	InterPro: IPR003587; Hedgehog_hlntn.		
DR	InterPro: IPR000320; HH_signal.		
DR	InterPro: IPR002203; Inteln.		
DR	InterPro: IPR001657; SonicHH.		
DR	Pfam: PF01085; HH_signal.		
DR	Pfam: PF01079; Hlntc_1.		
DR	PRINTS: PR00632; SONICHOG.		
DR	PRODOM: PD003042; HH_signal; 1.		
DR	SMART: SM00305; Hlntc_1.		
DR	SMART: SM00306; Hlntcn_1.		
SO	SEQUENCE 414 AA; 45945 MW; 50607BF3DB7C0DA3 CRC64;		

Query Match	67.68;	Score 1499;	DB 13;	Length 414;
Best Local Similarity	69.38;	Pred. No. 3e-115;		
Matches 289;	Conservative 52;	Mismatches 68;	Indels 8;	Gaps 5;

```

0Y 4 MLTRILLYVEFICALVSSLTJCGPGGJGRKHPPKLPYAKOIPVAEKTGASG 63
Db 1 MLMLTRILYAVICLSLSSVSGMGCGRGGRKHPPKLPPLAKOIPVAEKTGASG 60
0Y 64 RYEGKTIENSERFELTPNPNPDIIFKDEENTGADRLMTORCKOKJLALASVNMOPGV 123
Db 61 RYEGKTIENSERFELTPNPNYTDIIFDEENTGADRLMTORCKOKJLALASVNMOPGV 120
0Y 124 KLRVTEGMDDEGHHSEBSLHYEGRAVDITTSDDRKSXYGMRLARLAVAGFDWYTESKAH 183
Db 121 KLRVTEGMDDEGHHFEBSLHYEGRAVDITTSDDRKSXYGTLRLAVAGFDWYTESKAH 180
0Y 184 IHCSTKANSVAASGCGFPCSAVNHLEHGKTKYKPLSGDRLADADGRLLYSPEFL 243
Db 181 IHCSTKANSVAASGCGFPCSSVTYTDGKKFKVKALQTDGRLADADAGOPPYTDFEIM 240
0Y 244 FLDRRDSRSKLFYUJETROPARLLTFAAHLLFPAPOHNSOATGSGTSGOALFASNVKPG 303
Db 241 FIDDSOTIRLFYUETDSSG-KITLRNHLLEFG--HNSNTEBAHHGMSNVAFASQVRFC 297
0Y 304 QRVYVLGEGGQULLPASVHVSLEBRESGAYAPRLTAOGTILINFLVASCYAVIEEHSMAH 363
Db 298 QTVFVL--DAERLQPVYKRI-YTQEHGSGSPAPVTAOGTVVDDOVLASCYAVIODEHLAH 354
0Y 364 WAFAPFLAQOULLALCPDGAIPPLAATITTTGIIHHYSKLLRIGSWYLDGALPLRGM 420
Db 355 WALAPFLRAHAYSSILF--SSQOPASQKODSVHYSKLLYOLGWLMLDSHSIPLGM 409

```

RESULT 3			
057567			
ID	057567	PRELIMINARY;	PRT; 406 AA.
AC	057567;		
DT	01-JUN-1998	(TREMBLrel, 06, Created)	
DT	01-JUN-1998	(TREMBLrel, 06, Last sequence update)	
DT	01-JUN-2002	(TREMBLrel, 21, Last annotation update)	
DE	Hedgehog segment polarity homolog.		
OS	Notoptthalmus viridescens (Eastern newt) (Triturus viridescens)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi		
CC	Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae;		
CC	Notoptthalmus.		
OX	NCBI_Taxid=8316;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Stark D.R., Gates P.B., Brookes J.P., Ferretti P.;		
RT	"Hedgehog homologue from Notoptthalmus viridescens.";		
RL	Dev. Dyn. 0:0-0(1994).		
CC	-1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE		
CC	PHOSPHORIBOSYLTRANSFERASE FAMILY.		
DR	EMBL; AF047466; AAC03108.1; -.		
DR	HSSP; Q62226; 1VHM.		
DR	InterPro; IPR001167; Hedgehog_hint.		
DR	InterPro; IPR003586; Hedgehog_hntC.		
DR	InterPro; IPR003587; Hedgehog_hntN.		
DR	InterPro; IPR000320; HH_signal.		
DR	InterPro; IPR002375; Pr/py_trp_transf.		
DR	InterPro; IPR001657; Son1GH.		
DR	Pfam; PF01085; HH_signal; 1.		
DR	Pfam; PF01079; Hnt; 1.		
DR	PRINTS; PR00632; SON1GHOG.		
DR	ProDom; PD003042; HH_signal; 1.		
DR	SMART; SM00305; HntC; 1.		
DR	SMART; SM00306; HntN; 1.		
DR	PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.		
KW	Glycyltransferase; Transferase.		
SO	SEQUENCE 406 AA; 45072 MW; 56422CAB5314D4D0 CRC64;		





```

Db      64 GLENGRITDSERFHTLKNFNTDILFKDEKTAGDRFMTORCKDLNALAISVNMQWG 123
Oy      123 VKLRVTEGDEGHHSEELHEGRAVDITTSDBRSKTKMLARLAVEGDFWVYESNA 182
Db      124 VKLRVTEGDEGHHSEELHEGRAVDITTSDBRSKTKMLARLAVEGDFWVYESNA 183
Oy      183 HHCYSKAEENVAASGSGCFPSATVLEHGKTKLVKDLSPGDRVLADADGRLVSDPL 242
Db      184 HHCYSKAEENVAASGSGCFPSATVLEHGKTKLVKDLSPGDRVLADADGRLVSDPL 243
Oy      243 TELDRDSSRKLFYVETQPARALLTPAHLFVAPQHNOSATSTGQALFASNVNP 302
Db      244 TELDRDSSRKLFYVETQPARALLTPAHLFVAPQHNOSATSTGQALFASNVNP 298
Oy      303 GORVYVLEGGQOQLPASVSVSLREASGAYAPLTAOCTILINRYLASCAYIEESMA 362
Db      299 GEFLLTPDSDGGCFKRVKIVSYTMREE-KGATAPLTVHGTVVADVAMSCYALIESQALA 357
Oy      363 HMAFAPFLAQLALALCPDGAIPPT-AAVTTTGIMHYSRLYRIGSVLDGALHP 417
Db      358 HMAFAPFLAQLALALCPDGAIPPT-AAVTTTGIMHYSRLYRIGSVLDGALHP 410

```

## RESULT 6

```

ID      073803 PRELIMINARY: PRT: 442 AA.
AC      073803:
DT      01-AUG-1998 (TREMBLrel. 07, Created)
DT      01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DE      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
GN      Fugu hedgehog.
OS      Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC      Acanthomorphi; Acanthopterygii; Perciformes; Tetraodontiformes;
OC      Tetraodontidae; Takifugu.
OX      NCBI_TaxID=31033;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-99177347; PubMed-10077531;
RA      Gellner K., Brenner S.;
RT      "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu
RT      rubripes."
RL      Genome Res. 9:251-258(1999).
DR      EMBL: AF056116; AAC34384.1; -.
DR      HSP: O62226; 1VNH.
DR      InterPro: IPR001767; Hedgehog_hint.
DR      InterPro: IPR003586; Hedgehog_hintc.
DR      InterPro: IPR003587; Hedgehog_hintn.
DR      InterPro: IPR000320; HH_signal.
DR      InterPro: IPR002203; Inteln.
DR      Pfam: PF01085; HH_signal. 1.
DR      Pfam: PF01079; Hintc. 1.
DR      ProDom: PD003042; HH_signal. 1.
DR      SMART: SM00305; Hintc. 1.
DR      SMART: SM00306; Hintn. 1.
SQ      SEQUENCE 442 AA: 49286 MW: 1CB2BA23B4B748E8 CRC64:

```

```

Query Match      47.7%; Score 1059; DB 13; Length 442;
Best Local Similarity 50.5%; Pred. No. 6; Se-79;
Matches 223; Conservative 56; Mismatches 115; Indels 48; Gaps 9;

```

```

Oy      11 LVGFICALVSSGLTCGPGRGIGRRHPRKLTPLAYKQIFIPNAKTKGASGRYKKT 70
Db      14 LLAAMSCVWLVOG--CGGPGGIGITRPRKLTAMTYKQFPPLSENNGASGRAGKKT 70
Oy      71 RNSERKELTPYNPDIIFKDENTGADRLMTORCKDKLNALASVNMOPGVKLVTEG 130
Db      71 RNSERKELTPYNPDIIFKDENTGADRLMTORCKDKLNALASVNMOPGVKLVTEG 130
Oy      131 WEDGHHSEELHEGRAVDITTSDBRSKTKMLARLAVEGDFWVYESNAHHCYSKA 190

```

```

Db      131 WEDGHHSEELHEGRAVDITTSDBRSKTKMLARLAVEGDFWVYESNAHHCYSKA 190
Oy      191 ENSVAASGSGCFPSATVLEHGKTKLVKDLSPGDRVLADADGRLVSDPLFLDRMS 250
Db      191 DMSVAEKGSGCFPSATVLEHGKTKLVKDLSPGDRVLADADGRLVSDPLFLDRMS 250
Oy      251 SRKLFYVETQPARALLTPAHLFVAPQHNOSATSTGQALFASNVNP 307
Db      251 SRKLFYVETQPARALLTPAHLFVAPQHNOSATSTGQALFASNVNP 302
Oy      308 VLGGGQOQLPASVSVSLREASGAYAPLTAOCTILINRYLASCAYIEESMAHMA 367
Db      303 VHTAGG-EVHPKRVISIT-EESVGAVALPTEAGSVFVGVGLASSALVEDIQLAHMAG 360
Oy      368 PRL-----AQLALCPDGAIPPT-AAVTTTGIMHYSRLYRIGSVLDGALHP 395
Db      361 PRLSSVSQOLMAEPREESDGSKTPLOPHALVGRDKKCARNSTSVREAGPRGRTSEV 420
Oy      396 HMYSRLLYRIGSVLDGALHP 417
Db      421 HMYAQLHRLGWTVLNPDLEHP 442

```

## RESULT 7

```

ID      09GRAB PRELIMINARY: PRT: 447 AA.
AC      09GRAB:
DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      Hedgehog protein (Fragment).
OS      Gryllus bimaculatus (Two-spotted cricket).
OC      Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC      Gryllinae; Gryllus.
OX      NCBI_TaxID=6999;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-20461145; PubMed-11003837;
RA      Nish N., Inoue Y., Nozawa A., Saito M., Misumi Y., Ohuchi H.,
RA      Yoshioke H., Noji S.;
RT      "Correlation of diversity of leg morphology in Gryllus bimaculatus
RT      (cricket) with divergence in dpp expression pattern during leg
RT      development."
RL      Development 127:4373-4381(2000).
DR      EMBL: AB044709; BAB19658.1; -.
DR      HSP: O62226; 1VNH.
DR      InterPro: IPR001767; Hedgehog_hint.
DR      InterPro: IPR003586; Hedgehog_hintc.
DR      InterPro: IPR003587; Hedgehog_hintn.
DR      InterPro: IPR000320; HH_signal.
DR      InterPro: IPR001657; SonlchH.
DR      Pfam: PF01085; HH_signal. 1.
DR      Pfam: PF01079; Hintc. 1.
DR      PRINTS: PR00632; SONLCHOG.
DR      ProDom: PD003042; HH_signal. 1.
DR      SMART: SM00305; Hintc. 1.
DR      SMART: SM00306; Hintn. 1.
FT      NON-TER
SQ      SEQUENCE 447 AA: 48008 MW: 29AEFB061C3EB6F0 CRC64:

```

```

Query Match      47.1%; Score 1044.5; DB 5; Length 447;
Best Local Similarity 53.9%; Pred. No. 1e-77;
Matches 220; Conservative 48; Mismatches 123; Indels 17; Gaps 8;

```

```

Oy      3 EMLLTRILLVGFICALVSSGLTCGPGRGIGRRHPRKLTPLAYKQIFIPNAEKTGAS 62
Db      14 KMLLTRILLVSSGLTCGPGRGIGRRHPRKLTPLAYKQIFIPNAEKTGAS 70
Oy      63 GVEGKGTNSRFRKLTTPYNPDIIFKDENTGADRLMTORCKDKLNALASVNMQWG 122
Db      71 GAERVRVADDPFRDLVBNYNADIVFKDEGTGADRLMTORCKDKLNALASVNMQWG 130

```



[illegible]

ID	061676	PRELIMINARY;	PRT;	410 AA.
AC	061676:			
DT	01-AUG-1998	(TREMBlrel. 07, Created)		
DT	01-AUG-1998	(TREMBlrel. 07, Last sequence update)		
DT	01-JUN-2002	(TREMBlrel. 21, Last annotation update)		
DE	Hedgehog.			
OS	Lytechinus variegatus (Sea urchin).			
OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;			
OC	Echinoidea; Euechinoidea; Echinacea; Temnopneustidae;			
OC	Lytechinus.			
OX	NCHI_TaxID=7654;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hertzler P.L., McClay D.R.;			
RT	"A sea urchin hedgehog homolog."			
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF059606; AAC15065.1; -			
DR	HSSP; Q62226; 1VNH.			
DR	InterPro; IPR001677; Hedgehog_hnt.			
DR	InterPro; IPR003586; Hedgehog_hntC.			
DR	InterPro; IPR003587; Hedgehog_hntN.			
DR	InterPro; IPR003202; HH_signal.			
DR	InterPro; IPR002203; Intein.			
DR	InterPro; IPR001657; SonicH.			
DR	Pfam; PF01085; HH_signal; 1.			
DR	Pfam; PF01079; Hnt; 1.			
DR	PRINTS; PR00632; SONICHHOG.			
DR	ProDom; PD003042; HH_signal; 1.			
DR	SMART; SM00305; HntC; 1.			
DR	SMART; SM00306; HntN; 1.			
SO	SEQUENCE 410 AA; 46071 MW; 875C889DA6FDBED CRC64;			
Query Match	40.8%; Score 906; DB 5; Length 410;			
Best Local Similarity	48.1%; Pred. No. 2,4e-66;			
Matches 199;	Conservative 65; Mismatches 134; Indels 16; Gaps 10;			
OY	7 LTRILVGFICAL-LVSSGLTCGPGRGIGRRH-PKLTPLPAYKQFIPNVAEKTLAGSGR 64			
DB	6 MWKWTIVITPACIALITLTOACHPGRS-GKTSHPRRNRPIDQKQKVPNISDFTGASGP 64			
OY	YECKTTRSEPKELTPPNVNDPIFKDEENGCADLMTORCKDKLNALASVMNQPGVK 124			
DB	65 PEGRIDRDERSKSPNNNDIVKDEGTGADLMTORCKDKLNTLAISVMNEMPGK 124			
OY	LRVTEGMDEDGHHSSESLHYEGRAVDITTSDRDSKYGMLATLAVEAGFDWVYESKAI 184			
DB	125 LRVVAMDED-QPNVEPLHAEGRAVDITTSDDKDKKYGMLATLAVEAGFDWVYESKANV 183			
OY	185 HCSVAENSVAAKSGCGCPGSAIVLHREGTKLVKDLSPGDRVLADADGRLYSDFLTF 244			

OS Ambystoma mexicanum (Axolotl).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Caudata; Salamandroidae; Ambystomatidae;  
 OC Ambystoma.  
 OX NCBI\_TaxID=8296;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TOROK M.A., Izpienza-Belmonte J.C., Gardiner D.M., Bryant S.V.;  
 RL Submitted (OCY-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF031480; AAD18128.1; -  
 DR HSSP; 062226; LVHH.  
 DR InterPro; IPR001767; Hedgehog\_hnt.  
 DR InterPro; IPR000320; HH\_signal.  
 DR InterPro; IPR001657; SonichH.  
 DR Pfam; PF01085; HH\_signal; 1.  
 DR Pfam; PF01079; Hnt; 1.  
 DR PRINTS; PR00632; SONICHOG.  
 DR ProDom; PD003042; HH\_signal; 1.  
 FT NON\_TER 1  
 FT 150  
 SQ SEQUENCE 150 AA; 16599 MW; 986F65037A69A8AC CRC64;

## Query Match

Best Local Similarity 33.2%; Score 737; DB 13; Length 150;  
 Matches 140; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 86 DIFKDEMTGADRLMTGCKDKLNALAI SVNMQMPGVKLRVTEGWDGHHSESLHAYE 145  
 DB 1 DIFKDEMTGADRLMTGCKDKLNALAI SVNMQMPGVKLRVTEGWDGHHSESLHAYE 60  
 OY 146 GRAVDITTSDDRSKYGKMLARLAVENGFDWVYVESKAHHCYKAENSVANSGCCPEPS 205  
 DB 61 GRAVDITTSDDRSKYGKMLARLAVENGFDWVYVESKAHHCYKAENSVANSGCCPEPS 120  
 OY 206 ATVLEHGKTKLVKDLSPGDRVLADADGR 235  
 DB 121 AKVLEHGVTTPKDLRPGDRVLADADGR 150

## RESULT 11

O9XS16

ID O9XS16 PRELIMINARY; PRT; 139 AA.

AC O9XS16;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)  
 DE Sonic hedgehog (Fragment).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE="TOOTH GERM".  
 RA Koyama E., Iwamoto M., Ohmori T., Kurisu K., Wu C., Ookura T.,  
 RA Basalir M.M., Tucker T., Pacifici M.;  
 RT "Development of Stratum Intermedium and its Role as a Sonic Hedgehog-  
 RT Signaling Structure During Odontogenesis."  
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF144100; AAD33926.1; -  
 DR HSSP; 062226; LVHH.  
 DR InterPro; IPR000320; HH\_signal.  
 DR InterPro; IPR001657; SonichH.  
 DR Pfam; PF01085; HH\_signal; 1.  
 DR PRINTS; PR00632; SONICHOG.  
 DR ProDom; PD003042; HH\_signal; 1.  
 FT NON\_TER 1  
 FT 139  
 SQ SEQUENCE 139 AA; 15961 MW; C83B0B96982046B1 CRC64;

Query Match 33.1%; Score 735; DB 6; Length 139;  
 Best Local Similarity 99.3%; Pred. No. 5.8e-53;

Matches 138; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 46 AYKQFIPNVAEKTIGASGRYEGKITRNSERFKELTPNYNPDIIFKDEMTGADRLMTGCK 105  
 DB 1 AYKQFIPNVAEKTIGASGRYEGKITRNSERFKELTPNYNPDIIFKDEMTGADRLMTGCK 60  
 OY 106 KDKLNALAI SVNMQMPGVKLRVTEGWDGHHSESLHAYE GRAVDITTSDDRSKYGKMLA 165  
 DB 61 KDKLNALAI SVNMQMPGVKLRVTEGWDGHHSESLHAYE GRAVDITTSDDRSKYGKMLA 120  
 OY 166 RLAVENGFDWVYVESKAH 184  
 DB 121 RLAVENGFDWVYVESKAH 139

## RESULT 12

O9W6C1

ID O9W6C1 PRELIMINARY; PRT; 138 AA.

AC O9W6C1;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)  
 DE Sonic hedgehog protein (Fragment).  
 GN SHH.  
 OS Eleutherodactylus coqui (Puerto Rican coqui).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;  
 OC Eleutherodactylus.  
 OX NCBI\_TaxID=57060;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Carl T.E., Richardson M.K., Olsson L., Schlosser G., Klymkowsky M.W.,  
 RA Hanken J.;  
 RT "Differences in vertebrate limb development revealed by studies of the  
 RT direct developing frog E. coqui".  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF113403; AAD23436.1; -  
 DR HSSP; 062226; LVHH.  
 DR InterPro; IPR000320; HH\_signal.  
 DR InterPro; IPR001657; SonichH.  
 DR Pfam; PF01085; HH\_signal; 1.  
 DR PRINTS; PR00632; SONICHOG.  
 DR ProDom; PD003042; HH\_signal; 1.  
 FT NON\_TER 1  
 FT 138  
 SQ SEQUENCE 138 AA; 15751 MW; FFA156A17F4681F0 CRC64;

Query Match 32.1%; Score 713; DB 13; Length 138;  
 Best Local Similarity 97.1%; Pred. No. 3.8e-51;  
 Matches 134; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 51 IPNVAEKTIGASGRYEGKITRNSERFKELTPNYNPDIIFKDEMTGADRLMTGCKDKLN 110  
 DB 1 IPNVAEKTIGASGRYEGKITRNSERFKELTPNYNPDIIFKDEMTGADRLMTGCKDKLN 60  
 OY 111 ALAISVNMQMPGVKLRVTEGWDGHHSESLHAYE GRAVDITTSDDRSKYGKMLARLAVE 170  
 DB 61 ALAISVNMQMPGVKLRVTEGWDGHHSESLHAYE GRAVDITTSDDRSKYGKMLARLAVE 120  
 OY 171 AGFDWVYVESKAHHCY 188  
 DB 121 AGFDWVYVESKAHHCY 138

## RESULT 13

O9WV29

ID O9WV29 PRELIMINARY; PRT; 177 AA.

AC O9WV29;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)  
 DE Indian hedgehog protein (Fragment).  
 OS Rattus norvegicus (Rat).

```

OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RA Garges P.L., Meyer R.A. Jr., Brown C.A., Price D.K.;
RT "Indian hedgehog in rat.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF162914; AAD45372.1; -
DR HSSP: Q62226; 1YHH.
DR MEROPS: C46_003; -
DR InterPro: IPR001767; Hedgehog_hlnt.
DR InterPro: IPR003587; Hedgehog_hlntn.
DR InterPro: IPR003020; HH_signal.
DR InterPro: IPR001657; SonichH.
DR Pfam: PF01085; HH_signal; 1.
DR Pfam: PF01079; Hlnt; 1.
DR PRINTS: PRO0632; SONICHOG.
DR PRODOM: PD003042; HH_signal; 1.
DR SMART: SM00306; Hlntn; 1.
FT NON_TER 1
FT SEQUENCE 177 AA; 19739 MW; CA5871626A56565 CRC64;

Query Match
Best Local Similarity 31.4%; Score 697; DB 11; Length 177;
Matches 132; Conservative 14; Mismatches 31; Indels 0; Gaps 0;

Qy 100 LMTORCKDLNALAISVNONPQVKRLRTGEWDEGHHSESLHYEGAAVITTSRDRS 159
Db 1 LMTORCKDLRLSLAISVNONPQVKRLRTGEWDEGHHSESLHYEGAAVITTSRDRN 60
Qy 160 KYGMALARLAVAGEFDWVYSEKKAHICSVKAKENSVAAKSGCFPSATVHLEHGTRKYK 219
Db 61 KYGLLARLAVAGFDWVYSEKKAHICSVKAKENSVAAKSGCFPSAGAVHLETGERVALS 120
Qy 220 DLSFGDRLVLAADAGRLYSDFLFLDRMDSRKLEVIETRQPRARLLTAAHLLE 276
Db 121 AVKGDRLVLAAGEDGNPTFSVLLFLDRPRERLRAFYIETQDPRLALTPAHLLF 177

RESULT 14
ID 042234 PRELIMINARY; PRT; 137 AA.
AC 042234;
DT 01-JAN-1998 (TREMBLREL. 05, Created)
DT 01-JAN-1998 (TREMBLREL. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLREL. 20, Last annotation update)
DE Sonic hedgehog (fragment).
GN SHH.
OS Coccurnix coccurnix (common quail).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosaustra; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coccurnix.
OC NCBI_TaxID=9091;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=98167903; PubMed=9435297;
RA Borycki A.G., Mendham L., Emerson C.P. Jr.;
RT "Control of somite patterning by sonic hedgehog and its downstream
RT signal response genes.";
RL Development 125:777-790(1998).
DR EMBL: AF022882; AAB80949.1; -
DR MEROPS: C46_002; -
DR InterPro: IPR001767; Hedgehog_hlnt.
DR InterPro: IPR003586; Hedgehog_hlntc.
DR InterPro: IPR001657; SonichH.
DR Pfam: PF01079; Hlnt; 1.
DR PRINTS: PRO0632; SONICHOG.
DR SMART: SM00305; Hlntc; 1.
FT NON_TER 1
FT SEQUENCE 137 AA; 14571 MW; 7FD29DF815AF1532 CRC64;

```

```

Query Match
Best Local Similarity 30.7%; Score 682; DB 13; Length 137;
Matches 130; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 290 TSGCALPFSNVKPGQRYVVLCEGGQQLPASVSHSLREESGAVAPLTAOGTILINRYL 349
Db 2 TSGCAFFASNVKPGQRYVVLCEGGQQLPASVSHSLREESGAVAPLTAOGTILINRYL 61
Qy 350 ASCYAVIEHSMAMAFPRFLAOGILNALCPDGAIPATTTGIIHYSRLYRIGSV 409
Db 62 TSCTAVIEHSMAMAFPRFLAOGILNALCPDGAIPATTTGIIHYSRLYRIGSV 121
Qy 410 LDGDALHPLGNVAPAS 425
Db 122 LDGDALHPLGNVAPAS 137

RESULT 15
ID 096699 PRELIMINARY; PRT; 185 AA.
AC 096699;
DT 01-MAY-1999 (TREMBLREL. 10, Created)
DT 01-MAY-1999 (TREMBLREL. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLREL. 17, Last annotation update)
DE Hedgehog protein (fragment).
GN HH.
OS Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditytysia;
OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
RN 11
RP SEQUENCE FROM N.A.
RA Keys D.N., Lewis D.L., Selegue J.E., Pearson B.J., Goodrich L.V.,
RA Johnson R.L., Gates J., Scott M.P., Carroll S.B.;
RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
RT evolution.";
RL Science 0:0-0(1999).
DR EMBL: AF117742; AAD08931.1; -
DR HSSP: Q62226; 1YHH.
DR InterPro: IPR003020; HH_signal.
DR InterPro: IPR001657; SonichH.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PRO0632; SONICHOG.
DR PRODOM: PD003042; HH_signal; 1.
FT NON_TER 1
FT SEQUENCE 185 AA; 20745 MW; E9B9F4CE2FAAE662 CRC64;

Query Match
Best Local Similarity 28.1%; Score 624; DB 5; Length 185;
Matches 117; Conservative 21; Mismatches 30; Indels 2; Gaps 2;

Qy 61 ASGREYKTRNSERFKELTPYNPDILFKDEENTGADRLMTORCKDLNALAISVNONQ 120
Db 3 ASGPREGRTDRDEKFRDLVPYNPDIIDFKDEGTGADRLMTORCKELKNTLAISVNONQ 62
Qy 121 PGVKRLRTVEGDEGHHSESLHYEGRAVDITTSRDRSKYGMALARLAVAGEFDWVYTES 180
Db 63 PGVRLRTVEGDEGHHSESLHYEGRAVDITTSRDRSKYGMALARLAVAGEFDWVYEN 122
Qy 181 KAHICSVKAKENSVAAKSGCFPSATVHLEHGTRKYKDLSPGDRVLA 230
Db 123 RSYIHCSYKTESVGT-GAGCFPSGAVVHTENGPAD-TASLAKKGVLA 170

```

Search completed: February 20, 2003, 10:11:50  
 Job time : 30.3373 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:05:32 ; Search time 31.7046 Seconds  
(without alignments)  
1836.661 Million cell updates/sec

Title: US-09-827-110A-13

Perfect score: 2289  
Sequence: 1 MLLLRARCFVILASLLVC.....GTWLDSETHMIGMAVKSS 437

Scoring table:  
BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2289	100.0	437	16	AA077339	Mouse sonic hedgeh
2	2289	100.0	437	20	AA05857	Mouse sonic hedgeh
3	2289	100.0	437	20	AA05513	Mouse sonic hedgeh
4	2289	100.0	437	20	AA07768	Mouse sonic hedgeh
5	2289	100.0	437	20	AA09471	Mouse Shh hedgehog
6	2289	100.0	437	21	AA095975	Mouse sonic hedgeh
7	2289	100.0	437	21	AA095284	Mouse sonic hedgeh
8	2289	100.0	437	21	AA096246	Partial mouse Shh.
9	2289	100.0	437	21	AA070679	Mouse sonic hedgeh
10	2289	100.0	437	22	AA065746	Mouse sonic hedgeh

11	2289	100.0	437	22	AA085736	Mouse sonic hedgeh
12	2289	100.0	437	22	AA084672	Amino acid sequenc
13	2289	100.0	437	22	AA004685	Mouse sonic hedgeh
14	2289	100.0	437	22	AA05375	Mouse sonic hedgeh
15	2289	100.0	437	22	AA031220	Amino acid sequenc
16	2289	100.0	437	22	AA060263	Mouse sonic hedgeh
17	2289	100.0	437	23	AA099481	Mouse sonic hedgeh
18	2289	100.0	437	23	AA079135	Mouse sonic hedgeh
19	2289	100.0	437	23	AA020921	Mouse sonic hedgeh
20	2289	100.0	437	23	AA014295	Mouse Shh protein
21	2286	99.9	437	19	AA061488	Mouse sonic hedgeh
22	2286	99.9	437	22	AA050292	Mouse sonic hedgeh
23	2286	99.9	437	22	AA085086	Mouse sonic hedgeh
24	2286	99.9	437	22	AA097559	Mouse sonic hedgeh
25	2286	99.9	437	23	AA021264	437-mer wild-type
26	2286	99.9	437	23	AA019834	Mouse sonic hedgeh
27	2280	99.6	437	23	AA021266	Mouse sonic hedgeh
28	2277	99.5	437	23	AA021273	Sonic hedgehog pro
29	2277	99.5	437	23	AA021274	Sonic hedgehog pro
30	2276	99.4	437	23	AA021265	Sonic hedgehog pro
31	2272	99.3	437	23	AA021272	Sonic hedgehog pro
32	2270	99.2	437	23	AA021267	Sonic hedgehog pro
33	2266	98.0	437	23	AA021271	Sonic hedgehog pro
34	2263	98.9	437	23	AA021269	Sonic hedgehog pro
35	2263	98.5	437	23	AA021274	Sonic hedgehog pro
36	2237	97.7	437	23	AA021268	Sonic hedgehog pro
37	2161	94.4	437	16	AA080071	Rat Vhh-1
38	2065	90.2	462	22	AA085085	Human sonic hedgeh
39	2065	90.2	463	22	AA085738	Human sonic hedgeh
40	2065	90.2	475	16	AA077341	Human sonic hedgeh
41	2065	90.2	475	20	AA050589	Human sonic hedgeh
42	2065	90.2	475	20	AA050515	Human sonic hedgeh
43	2065	90.2	475	20	AA097770	Human sonic hedgeh
44	2065	90.2	475	20	AA094473	Human Shh hedgehog
45	2065	90.2	475	21	AA095977	Human Sonic hedgeh

## ALIGNMENTS

RESULT 1	
AA077339	
ID	AA077339 standard; Protein; 437 AA.
AC	AA077339;
DT	08-MAR-1996 (first entry)
DE	Mouse sonic hedgehog protein.
OS	Mouse; sonic hedgehog protein; probe; primer; diagnostic;
XX	nervous system disorder; gene therapy; antibody.
XX	Mus musculus.
XX	Key
XX	Peptide
XX	Location/Qualifiers
XX	1..24
XX	/note= "signal peptide"
XX	25..30
XX	/note= "conserved sequence (AA077349)"
XX	279
XX	/note= "N-linked glycosylation site"
XX	W09518856-A1.
XX	13-JUL-1995.
XX	30-DEC-1994;
XX	94MO-US14992.
XX	14-DEC-1994;
XX	94US-0356060.
XX	30-DEC-1993;
XX	93US-0176427.
XX	(HARD ) HARVARD COLLEGE.

PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.  
XX Ingham PW, McMahon AP, Tabin CJ;  
XX WPI: 1995-255060/33.  
DR N-PSDB; AAO91637.  
XX Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful  
PT to treat degenerative nervous system disorder(s) and in gene  
PT therapy.  
XX Claim 17; Page 139-41; 210pp; English.  
XX The sequence represents a mouse sonic hedgehog protein,  
CC homologous to a Drosophila hedgehog protein (AA077337), and is  
CC encoded by a cDNA isolated from an 8.5-day post coitum mouse cDNA  
CC library. Probes and primers derived from the sonic hedgehog gene  
CC may be used as diagnostic agents for neuromuscular, autonomic or  
CC central nervous system disorders, and the gene may also be used in  
CC gene therapy. Antibodies generated from the protein may be used  
CC as therapeutic or research reagents.

XX Sequence 437 AA:

Query Match 100.0%; Score 2289; DB 16; Length 437;  
Best Local Similarity 100.0%; Pred. No. 9,4e-226;  
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLLLARCFVIILASSLLVCPGLACGPGRGFRKRRHKLTPLAYKQFINVAEKTIGAS 60  
DB 1 MLLLLARCFVIILASSLLVCPGLACGPGRGFRKRRHKLTPLAYKQFINVAEKTIGAS 60  
OY 61 GREGKITSRSEKFKELTPYNYNDIIFKDEENTGADRLMTQRCCKDKLNALASVMNQMPG 120  
DB 61 GREGKITSRSEKFKELTPYNYNDIIFKDEENTGADRLMTQRCCKDKLNALASVMNQMPG 120  
OY 121 VRLRVTEGMDGDHSEESIHYEGRAVDITTSRDNSKYGMALARAVEAGFDWVYVESKA 180  
DB 121 VRLRVTEGMDGDHSEESIHYEGRAVDITTSRDNSKYGMALARAVEAGFDWVYVESKA 180  
OY 181 HHCSSVAENSVAAGSGCGPGSATVHLDEGGTKLVLDLRPGDVLAAADGGRLYSDFL 240  
DB 181 HHCSSVAENSVAAGSGCGPGSATVHLDEGGTKLVLDLRPGDVLAAADGGRLYSDFL 240  
OY 241 TFLDRDEGAARVYVETLEPRERLLTAAHLLEFVAPHNDSGPTPGSALFASVRPQOR 300  
DB 241 TFLDRDEGAARVYVETLEPRERLLTAAHLLEFVAPHNDSGPTPGSALFASVRPQOR 300  
OY 301 VYVVAERGGRRLLPRAVHVTLEERAGAVAPLTAGTILINRYLASCYAVEEHSMAH 360  
DB 301 VYVVAERGGRRLLPRAVHVTLEERAGAVAPLTAGTILINRYLASCYAVEEHSMAH 360  
OY 361 RAFAFRLLAALLAALAPARTDGGGGSIIPAOSATEARGAEPAGIHWYSOLLHYHGTW 420  
DB 361 RAFAFRLLAALLAALAPARTDGGGGSIIPAOSATEARGAEPAGIHWYSOLLHYHGTW 420  
OY 421 LLDSETHPLGMAVKSS 437  
DB 421 LLDSETHPLGMAVKSS 437

RESULT 2

AA05857 standard; Protein: 437 AA.

AA05857;

02-AUG-1999 (first entry)

Mouse Sonic hedgehog Shh protein.

Sonic hedgehog; Shh; mouse; epithelial tissue; epithelium;

cutaneous tissue; skin; hair; wound healing; vulvectomy;

KW burn; skin grafting; pressure sore; ulcer; ulcerative colitis;  
KW alopecia; psoriasis; keratosis; acne; comedogenic lesion;  
KW folliculitis; pseudofolliculitis; keratoacanthoma; callosities;  
KW Darier's disease; scar; autoimmune disease; pemphigus;  
KW epidermolysis; lupus lesion; desquamative lesion; carcinoma;  
KW therapy; hedgehog therapeutic; ptc therapeutic; patched.  
XX Mus musculus.  
XX WO920298-A1.  
XX 29-APR-1999.  
XX 20-OCT-1998; 98WO-US22227.  
XX 11-SEP-1998; 98US-0151999.  
XX 20-OCT-1997; 97US-0955552.  
XX (ONTO-) ONTOGENY INC.  
XX Wang EA;  
XX WPI: 1999-288170/24.  
XX N-PSDB; AAX25620.  
XX Use of hedgehog polypeptides on patched therapeutics  
PS Claim 26; Page 124-125; 146pp; English.

XX The present sequence represents mouse Sonic hedgehog protein  
CC Shh. The invention relates to a method for modulating the growth  
CC state an epithelial cell by ecotopically contacting the epithelial  
CC cell, in vitro or in vivo, with a hedgehog therapeutic (i.e. a  
CC hedgehog polypeptide or gene therapy construct) or ptc therapeutic  
CC (i.e. a small organic molecule that mimics the effect of hedgehog  
CC proteins on patched signaling, or activates or potentiates patched  
CC signaling) in an amount effective to alter the rate of proliferation  
CC of the epithelial cell. The hedgehog therapeutic preferably  
CC comprises at least a bioactive extracellular portion of a hedgehog  
CC protein (see AA05854-62) encoded by a vertebrate hedgehog gene (see  
CC AAX25617-25), especially a human hedgehog gene. Promotion of  
CC proliferation of epithelial cells can be used to control a wound  
CC healing process in e.g. burn treatment, skin regeneration, skin  
CC grafting, pressure sore treatment, dermal ulcer treatment, post  
CC surgery scar reduction or treatment of ulcerative colitis (claimed).  
CC It can also be used to induce hair growth for the treatment of  
CC alopecia (claimed). Inhibition of the growth of epithelial tissue  
CC can be used to treat or prevent hyperplastic or neoplastic  
CC conditions, e.g. psoriasis, keratosis, acne, comedogenic lesions,  
CC folliculitis and pseudofolliculitis, keratoacanthoma, callosities,  
CC Darier's disease, keloids, hypertrophic scars, or autoimmune  
CC disorders, e.g. aphthous ulcers, pemphigus vulgaris, pemphigus  
CC foliaceus, pemphigus vegetans, pemphigus erythematous,  
CC epidermolysis, lupus lesions, desquamative lesions or carcinomas.  
CC The methods can also be used to counteract the effects of ageing on  
CC skin.

XX Sequence 437 AA:

Query Match 100.0%; Score 2289; DB 20; Length 437;  
Best Local Similarity 100.0%; Pred. No. 9,4e-226;  
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLLLARCFVIILASSLLVCPGLACGPGRGFRKRRHKLTPLAYKQFINVAEKTIGAS 60  
DB 1 MLLLLARCFVIILASSLLVCPGLACGPGRGFRKRRHKLTPLAYKQFINVAEKTIGAS 60  
OY 61 GREGKITSRSEKFKELTPYNYNDIIFKDEENTGADRLMTQRCCKDKLNALASVMNQMPG 120  
DB 61 GREGKITSRSEKFKELTPYNYNDIIFKDEENTGADRLMTQRCCKDKLNALASVMNQMPG 120  
OY 121 VRLRVTEGMDGDHSEESIHYEGRAVDITTSRDNSKYGMALARAVEAGFDWVYVESKA 180  
DB 121 VRLRVTEGMDGDHSEESIHYEGRAVDITTSRDNSKYGMALARAVEAGFDWVYVESKA 180

DB 121 VRLRVTEGMDGDGHSESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWVYTESKA 180  
QY 181 HHCYSVAENSVAAKSGCGPGSATVHLDEGGTKLVKDLRPGDRVLAADQGRLLYSDFL 240  
DB 181 HHCYSVAENSVAAKSGCGPGSATVHLDEGGTKLVKDLRPGDRVLAADQGRLLYSDFL 240  
QY 241 TFLDRDGGAKKVFYVETLEPRERLLTAAHLFPVAPHNDSGPTPGPSALFASVRPGR 300  
DB 241 TFLDRDGGAKKVFYVETLEPRERLLTAAHLFPVAPHNDSGPTPGPSALFASVRPGR 300  
QY 301 VYVAERGGDRLLPAVHVSVTLEEEAGAVAPLTAGTLLINRVLASCAVAIEESHMAH 360  
DB 301 VYVAERGGDRLLPAVHVSVTLEEEAGAVAPLTAGTLLINRVLASCAVAIEESHMAH 360  
QY 361 RAFAFRLAHALLAALAPARTDGGGGGSIIPAOSATFARGAEPAGIHWYSOLLXHGTM 420  
DB 361 RAFAFRLAHALLAALAPARTDGGGGGSIIPAOSATFARGAEPAGIHWYSOLLXHGTM 420  
QY 421 LLDSETHMPLGMAVKSS 437  
DB 421 LLDSETHMPLGMAVKSS 437

RESULT 3  
AA05513  
ID AAY05513 standard; Protein; 437 AA.  
AC AAY05513;  
XX  
DT 05-JUL-1999 (first entry)  
XX  
DE Mouse Sonic hedgehog protein Shh.  
XX  
KM Sonic hedgehog; Shh protein; mouse; hedgehog therapeutic;  
KW plc therapeutic; patched; signal transduction; muscle atrophy;  
XX cachexia; muscular myopathy; myoblastic sarcoma; therapy.  
XX  
OS Mus sp.  
XX  
PN W09910004-A2.  
XX  
PD 04-MAR-1999.  
XX  
PE 28-AUG-1998; 98WO-US19722.  
XX  
PR 29-AUG-1997; 97US-0057394.  
XX  
PA (ONTO-) ONTOGENY INC.  
XX  
PI Bladgen CS, Currie PD, Hughes SM, Ingham PW;  
XX  
DR WPI; 1999-243557/20.  
XX  
PT N-PSDB; AAX25101.  
XX  
PT A new method to regulate muscle growth  
XX  
PS  
XX  
XX Disclosure; Page 115-116; 130pp; English.  
XX  
XX The present sequence is mouse Sonic hedgehog protein Shh. The  
XX invention relates to a method for modulating the formation and/or  
XX maintenance of muscle tissue by ecotopically contacting muscle  
XX cells, especially muscle stem/progenitor cells, in vitro or in  
XX vivo, with a hedgehog therapeutic (i.e. hedgehog polypeptides and  
XX gene therapy constructs) or plc therapeutic (i.e. a small organic  
XX molecule that mimics the effect of hedgehog proteins on patched  
XX signalling, or activates or potentiates patched signalling) in an  
XX amount effective to alter the growth state of the treated cells.  
XX Also claimed is a method for treatment or prevention of disorders  
XX of, or surgical or cosmetic repair of, such muscle tissues, by  
XX administering a hedgehog polypeptide or plc therapeutic. The  
XX disorder may be muscle atrophy, in particular skeletal muscle  
XX atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy  
XX (all claimed). The hedgehog polypeptide or plc therapeutic can

CC Inhibit growth of myoblastic-derived tissue to provide treatment of  
CC hyperplastic or neoplastic growth of muscle tissue such as in  
CC myoblastic sarcoma (also claimed). The hedgehog therapeutic  
CC preferably comprises at least a bioactive extracellular portion of  
CC a hedgehog protein (see AAY05510-19) encoded by a vertebrate hedgehog  
CC gene (see AAX25098-107), especially a human hedgehog gene.  
XX  
SQ Sequence 437 AA:  
Query Match 100.0%; Score 2289; DB 20; Length 437;  
Best Local Similarity 100.0%; Pred. No. 9,4e-226;  
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLLILACFLVILASSILVCPGLACGPGRGFRKRRHKLTPLAYKQFINVAEKTIGAS 60  
DB 1 MLLILACFLVILASSILVCPGLACGPGRGFRKRRHKLTPLAYKQFINVAEKTIGAS 60  
QY 61 GRYEGKITRNSERKELTPYVNDIIFKDEENTGADRLMTORCKDKLINALAISVMQWPG 120  
DB 61 GRYEGKITRNSERKELTPYVNDIIFKDEENTGADRLMTORCKDKLINALAISVMQWPG 120  
QY 121 VRLRVTEGMDGDGHSESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWVYTESKA 180  
DB 121 VRLRVTEGMDGDGHSESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWVYTESKA 180  
QY 181 HHCYSVAENSVAAKSGCGPGSATVHLDEGGTKLVKDLRPGDRVLAADQGRLLYSDFL 240  
DB 181 HHCYSVAENSVAAKSGCGPGSATVHLDEGGTKLVKDLRPGDRVLAADQGRLLYSDFL 240  
QY 241 TFLDRDGGAKKVFYVETLEPRERLLTAAHLFPVAPHNDSGPTPGPSALFASVRPGR 300  
DB 241 TFLDRDGGAKKVFYVETLEPRERLLTAAHLFPVAPHNDSGPTPGPSALFASVRPGR 300  
QY 301 VYVAERGGDRLLPAVHVSVTLEEEAGAVAPLTAGTLLINRVLASCAVAIEESHMAH 360  
DB 301 VYVAERGGDRLLPAVHVSVTLEEEAGAVAPLTAGTLLINRVLASCAVAIEESHMAH 360  
QY 361 RAFAFRLAHALLAALAPARTDGGGGGSIIPAOSATFARGAEPAGIHWYSOLLXHGTM 420  
DB 361 RAFAFRLAHALLAALAPARTDGGGGGSIIPAOSATFARGAEPAGIHWYSOLLXHGTM 420  
QY 421 LLDSETHMPLGMAVKSS 437  
DB 421 LLDSETHMPLGMAVKSS 437

RESULT 4  
AAW97768  
ID AAW97768 standard; Protein; 437 AA.  
AC AAW97768;  
XX  
DT 21-MAY-1999 (first entry)  
XX  
DE Mouse Sonic hedgehog (Shh) protein.  
XX  
KM Sonic hedgehog; Shh protein; mouse; dopaminergic; GABA-ergic;  
KW plc therapeutic; patched; signal transduction; Parkinson's disease;  
XX Huntington's disease; amyotrophic lateral sclerosis;  
XX cerebral ischaemia; hypoxia; neuroprotective; therapy.  
XX  
OS Mus sp.  
XX  
PN W09904775-A2.  
XX  
PD 04-FEB-1999.  
XX  
PE 24-JUL-1998; 98WO-US15419.  
XX  
PR 24-JUL-1997; 97US-0900220.  
XX  
PA (ONTO-) ONTOGENY INC.

PI Mahanthappa NK, Miao N, Pang K, Wang M;  
 XX WPI; 1999-142578/12.  
 DR N-PSDB; AAX07274.  
 XX  
 PT Increasing the survival of neuronal, dopaminergic and GABA-nergic  
 PT cells - by using a ptc therapeutic such as a protein kinase  
 PT inhibitor, or an agent derived from hedgehog polypeptides, useful in  
 PT the treatment of Parkinson's disease  
 XX  
 PS Disclosure; Page 89-91; 138pp; English.  
 XX  
 CC This polypeptide is mouse Shh Sonic hedgehog protein. The  
 CC invention is based on the finding that hedgehog proteins are useful  
 CC as protective agents in the treatment and prophylaxis of  
 CC neurodegenerative disorders resulting from the loss of dopaminergic  
 CC and/or GABA-nergic neurons, or the general loss of tissue from the  
 CC substantia nigra. Exemplary disorders include Parkinson's disease,  
 CC Huntington's disease (both claimed), amyotrophic lateral sclerosis  
 CC and cerebral ischaemia. The invention relates to hedgehog  
 CC therapeutics (i.e. hedgehog polypeptides and gene therapy constructs  
 CC e.g. constructs encoding recombinant hedgehog polypeptides and  
 CC trans-activation constructs for altering hedgehog gene regulatory  
 CC sequences) and ptc therapeutics (i.e. agents which mimic the effect  
 CC of naturally occurring hedgehog proteins on patched signalling)  
 CC that are effective in both human and animal subjects. Human ihh  
 CC and dhb polypeptides (see AAW97763-64) are preferred. The products  
 CC can also be used for the maintenance of differentiated neurons in  
 CC cultures, and to enhance the implantation of such neuronal cells in  
 CC an animal. They can be used to prevent or treat neurodegenerative  
 CC conditions arising from the use of certain drugs, and in the  
 CC prevention and/or treatment of hypoxia, e.g. as a neuroprotective  
 CC agent.  
 XX  
 SQ Sequence 437 AA;  
 Query Match 100.0%; Score 2289; DB 20; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 9.4e-226;  
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLLILARCFVLIASSLVCPGLACGPGRGKRRHPKRLPLAYKQIPNVAEKTGAS 60  
 DB 1 MLLILARCFVLIASSLVCPGLACGPGRGKRRHPKRLPLAYKQIPNVAEKTGAS 60  
 QY 61 GREGKTRNSERKELTPYNDPIIFKDEENTGADRLMTORCKDKLNALAIISVNMWPG 120  
 DB 61 GREGKTRNSERKELTPYNDPIIFKDEENTGADRLMTORCKDKLNALAIISVNMWPG 120  
 QY 121 VRLVTEGMDDEGHSESLHYEGRAVDITTSRDSKYGKMLARLAVEAGFDWYTESKA 180  
 DB 121 VRLVTEGMDDEGHSESLHYEGRAVDITTSRDSKYGKMLARLAVEAGFDWYTESKA 180  
 QY 181 HHCYSKAENSVAAKSGCGPGSATVHLEGGCKLVKDLRPGDRLVLAADOGGLXSDFL 240  
 DB 181 HHCYSKAENSVAAKSGCGPGSATVHLEGGCKLVKDLRPGDRLVLAADOGGLXSDFL 240  
 QY 241 TFLDRREGAKKYYETETLEPRRLTLTAHLLFVAHPNDSCPTPGSALFASRVAPGQR 300  
 DB 241 TFLDRREGAKKYYETETLEPRRLTLTAHLLFVAHPNDSCPTPGSALFASRVAPGQR 300  
 QY 301 VYVVAERGGDRRLPAAVHSVTLREEAGAVAPLTAHGILLNRLVASCYAVIEEHSMAH 360  
 DB 301 VYVVAERGGDRRLPAAVHSVTLREEAGAVAPLTAHGILLNRLVASCYAVIEEHSMAH 360  
 QY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATEARGAEPATGIMWYSOLLTHIGTW 420  
 DB 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATEARGAEPATGIMWYSOLLTHIGTW 420  
 QY 421 LLDSTHAPLGMAYKSS 437  
 DB 421 LLDSTHAPLGMAYKSS 437  
 QY 421 LLDSTHAPLGMAYKSS 437  
 DB 421 LLDSTHAPLGMAYKSS 437

RESULT 5  
 AAW94471  
 ID AAW94471 standard; Protein; 437 AA.  
 XX  
 AC AAW94471;  
 XX  
 DT 29-APR-1999 (first entry)  
 XX  
 DE Mouse Shh hedgehog protein sequence.  
 XX  
 KW Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;  
 KW brain infarction; cerebral infarction; transient ischemic attack;  
 KW stroke; cerebral infarct volume; spinal cord; oedema; trauma;  
 KW haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.  
 OS Mus sp.  
 XX  
 PN WO9900117-A2.  
 XX  
 PD 07-JAN-1999.  
 XX  
 PF 26-JUN-1998; 98WO-US13387.  
 XX  
 PR 27-JUN-1997; 97US-0883656.  
 XX  
 PA (ONTO-) ONTOGENY INC.  
 XX  
 PI Mahanthappa NK;  
 XX  
 DR WPI; 1999-095458/08.  
 DR N-PSDB; AAX16185.  
 XX  
 PT Method for limiting damage to neurons caused by ischemic or epoxic  
 PT conditions - is used for the treatment and prevention of e.g.  
 PT cerebral infarction, stroke and transient ischemic attacks  
 XX  
 PS Disclosure; Page 68-70; 104pp; English.  
 XX  
 CC A method has been developed for limiting the damage to neuronal cells by  
 CC ischemic or epoxic conditions by administering a ptc (patched)  
 CC therapeutic agent to reduce cerebral infarct volume (CIIV). Damage to  
 CC neuronal cells can also be limited by administering a gene activation  
 CC construct which recombines with the genomic hedgehog gene to provide a  
 CC heterologous transcription regulator linked to the coding region of this  
 CC gene. Administration of the ptc therapeutic agent is used to protect  
 CC cerebral tissues against ischemic injury; to treat cerebral infarct or  
 CC ischemia; stroke (thrombotic or embolic) and transient ischemic  
 CC attacks. It may also be used as a prophylactic in many other cases of  
 CC injury to the brain or spinal cord, oedema caused by trauma, haemorrhage  
 CC and encephalomyelitis, or in conjunction with (coronary bypass) surgery.  
 CC treatment (which may be prophylactic) is used where ischemic/epoxic  
 CC conditions may cause cerebral hypoxia, or progressive loss of neurons  
 CC due to oxygen depletion, including in patients with hypotension. The  
 CC treatment reduces CIIV by at least 25, particularly at least 70%. The  
 CC present sequence represents a hedgehog sequence given in the present  
 CC invention.  
 CC  
 SQ Sequence 437 AA;  
 Query Match 100.0%; Score 2289; DB 20; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 9.4e-226;  
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLLILARCFVLIASSLVCPGLACGPGRGKRRHPKRLPLAYKQIPNVAEKTGAS 60  
 DB 1 MLLILARCFVLIASSLVCPGLACGPGRGKRRHPKRLPLAYKQIPNVAEKTGAS 60  
 QY 61 GREGKTRNSERKELTPYNDPIIFKDEENTGADRLMTORCKDKLNALAIISVNMWPG 120  
 DB 61 GREGKTRNSERKELTPYNDPIIFKDEENTGADRLMTORCKDKLNALAIISVNMWPG 120  
 QY 121 VRLVTEGMDDEGHSESLHYEGRAVDITTSRDSKYGKMLARLAVEAGFDWYTESKA 180  
 DB 121 VRLVTEGMDDEGHSESLHYEGRAVDITTSRDSKYGKMLARLAVEAGFDWYTESKA 180





KW Sonic hedgehog; Shh; mouse; excitotoxicity; Parkinson's disease;  
 KW Huntington's disease; neuronal degeneration; neuroprotective;  
 KW dopaminergic; GABAergic; substantia nigra; therapy.  
 OS Mus musculus.  
 PN MO200035948-A1.  
 XX 22-JUN-2000.  
 PD 03-DEC-1999; 99WO-US28721.  
 PF 03-DEC-1999; 99WO-US25676.  
 PR 03-DEC-1998; 98WO-US25676.  
 PR 27-JAN-1999; 99US-0238243.  
 PR 03-JUN-1999; 99US-0325602.  
 PA (BIOJ) BIOGEN INC.  
 PA (ONTO-) ONTOGENY INC.  
 PI Galdes A, Mahanthappa N;  
 PT WPI: 2000-431570/37.  
 DR N-PSDB: AAA27879.  
 XX Treating disorders involving exotoxicity, e.g. trauma, hypoglycemia,  
 PT senile dementia and Korsakoff's disease, by using lipophilic modified  
 PT hedgehog polypeptide -  
 PS Disclosure; Page 154-155; 174pp: English.  
 XX The present sequence of that of mouse Sonic hedgehog (Shh)  
 CC protein. The invention relates to a method for promoting the  
 CC survival and/or functional performance of neuronal cells,  
 CC especially substantia nigra, dopaminergic or GABAergic neurons that  
 CC are susceptible to exotoxicity, by contacting the cells, in vitro  
 CC or in vivo, with a lipophilic (e.g. cholesterol) modified hedgehog  
 CC polypeptide. The method is used to treat or prevent Parkinson's  
 CC disease, Huntington's disease, domoic acid poisoning, spinal cord  
 CC trauma, hypoglycemia, mechanical trauma to the nervous system,  
 CC senile dementia, Korsakoff's disease, schizophrenia, AIDS dementia,  
 CC multi-infarct dementia, mood disorders, depression, chemical  
 CC toxicity, neuronal damage associated with uncontrolled seizures  
 CC such as epileptic seizures, neuronal injury associated with HIV and  
 CC AIDS, neurodegeneration associated with Down's syndrome,  
 CC neuropathic pain syndrome, olivopontocerebral atrophy, amyotrophic  
 CC lateral sclerosis, mitochondrial abnormalities, Alzheimer's disease,  
 CC hepatic encephalopathy, Tourette's syndrome and drug addiction (all  
 CC claimed). The lipophilic modified hedgehog polypeptide is also  
 CC useful for promoting survival and/or functional performance of  
 CC neuronal cells susceptible to exotoxicity.  
 XX  
 SQ Sequence 437 AA:  
 Query Match 100.0%; Score 2289; DB 21; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-226;  
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 TFLDRDEGAKVYVYETLEPERELLTPAHLLEFVAPHNDSGPTPGPSALEFASRVRCOR 300  
 DB 241 TFLDRDEGAKVYVYETLEPERELLTPAHLLEFVAPHNDSGPTPGPSALFASRVRCOR 300  
 QY 301 VYVAERGGDRRLPAAVHSTLREEGAGVAPLTAGTTLINRLASCVAVIEHSNAH 360  
 DB 301 VYVAERGGDRRLPAAVHSTLREEGAGVAPLTAGTTLINRLASCVAVIEHSNAH 360  
 QY 361 RAFAPEFLAHLAALAPARTDGGGGSIPAAQSTEARGAEPYAGIHMYSQLYHIGTW 420  
 DB 361 RAFAPEFLAHLAALAPARTDGGGGSIPAAQSTEARGAEPYAGIHMYSQLYHIGTW 420  
 QY 421 LLDSETHPLGMAYKSS 437  
 DB 421 LLDSETHPLGMAYKSS 437  
 RESULT 8  
 ID AAY96246 standard; Protein; 437 AA.  
 XX AAY96246;  
 AC AAY96246;  
 XX 11-SEP-2000 (first entry)  
 DE Partial mouse Shh.  
 XX Mouse; sonic hedgehog; Shh; neuromuscular disorder; neuropathy;  
 KW Guillain-Barre syndrome; peripheral neuropathy; diabetes; alcoholism;  
 KW chronic inflammatory demyelinating polyneuropathy; CIPD;  
 KW gene therapy; infection; inflammation; hereditary neuropathy;  
 KW Charcot-Marie-Tooth disease; vasculitis; lung cancer; tumour;  
 KW multiple myeloma; nutritional imbalance; kidney disease;  
 KW hypothyroid neuropathy; trauma; Refsum's disease; Abetalipoproteinemia;  
 KW Tangier disease; Krabbe's disease; Metachromatic leukodystrophy;  
 KW Fabry's disease; CMT; GBS; Dejerine-Sottas syndrome; acute neuropathy;  
 KW amyotrophic lateral sclerosis; ALS; Miller-Fisher syndrome; amyloidosis;  
 KW hereditary sensory neuropathy type II; HSN II; B-cell lymphoma;  
 KW Waldenstrom's Macroglobulinaemia; Chronic Lymphocytic Leukaemia;  
 KW neuroprotective; cytoprotective; patched-mediated signal transduction.  
 XX  
 OS Mus sp.  
 XX WO200027422-A2.  
 PN 18-MAY-2000.  
 PD 08-NOV-1999; 99WO-US26334.  
 PF 08-NOV-1999; 98US-0187387.  
 PR 06-NOV-1998;  
 XX (BIOJ) BIOGEN INC.  
 PA (ONTO-) ONTOGENY INC.  
 PI Galdes A, Mahanthappa N;  
 PT WPI: 2000-387341/33.  
 DR N-PSDB: AAA30277.  
 XX Novel method of preventing deterioration of peripheral nerves, useful  
 PT for treating or preventing neuropathy, e.g. where associated with  
 PT diabetes or viral infection, by administering hedgehog or patched agent  
 PT -  
 XX Claim 7: Page 131-132; 152pp: English.  
 PS The present sequence is the partial mouse sonic hedgehog protein, Shh.  
 CC This sequence inhibits expression of the patched gene which has been  
 CC implicated in neuromuscular disorders (neuropathies). This sequence may  
 CC therefore be used for treating neuromuscular disorders i.e. preventing  
 CC degeneration in function of motor or sensory nerves and protecting  
 CC peripheral nerve cells under conditions that normally cause neuropathy.  
 CC A variety of neuromuscular disorders may be treated: Guillain-Barre



Db 361 RAFAFRLAHALLAALAPARTDGGGGSIIPAQSATEARGAEPAGIHWYSQLYHIGTW 420  
 QY 421 LLDSETHMPLGMAVKSS 437  
 Db 421 LLDSETHMPLGMAVKSS 437

RESULT 10  
 AAG65746  
 ID AAG65746 standard; Protein: 437 AA.  
 AC AAG65746;  
 XX  
 XX  
 DT 07-JAN-2002 (first entry)  
 XX  
 DE Mouse sonic hedgehog (Shh) polypeptide.  
 XX  
 KW Adipocyte; hedgehog polypeptide; desert hedgehog; indian hedgehog; Dhh;  
 XX Ihh; sonic hedgehog; Shh; therapeutic; cytostatic; tumor.  
 OS Mus sp.  
 XX  
 PN WO200164238-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 28-FEB-2001; 2001WO-US06450.  
 XX  
 PR 29-FEB-2000; 2000US-186058P.  
 XX  
 PA (CURT-) CURIS INC.  
 XX  
 PI Zehentner B, Leser-Reiff U, Bartscher H;  
 XX WPI: 2001-607352/69.  
 DR N-PSDB; AA166774.  
 XX

PT Method for regulating formation and/or maintenance of adipocyte tissue  
 by contacting pre-adipocyte or adipocyte cells with a hedgehog  
 polypeptide or ptc therapeutic

PS Disclosure: Page 97-99; 132pp; English.

CC The invention provides a method for regulating formation and/or  
 CC maintenance of adipocyte tissue that comprises contacting pre adipocyte  
 CC or adipocyte cells with a hedgehog polypeptide or ptc therapeutic. The  
 CC method is used for regulating the growth state of an adipocyte stem/  
 CC progenitor cell, and treating or preventing disorders of, or surgical or  
 CC cosmetic repair of, adipocyte tissues, e.g. for treating or preventing  
 CC hyperplastic or neoplastic conditions affecting adipocyte tissue, such  
 CC as soft tissue tumors, especially adipose cell tumors, e.g. lipomas,  
 CC fibrolipomas, lipoblastomas, lipomatosis, hibernomas, hemangiomas and/or  
 CC liposarcomas. Hedgehog polypeptides can be used in combination with other  
 CC therapeutic agents. The present sequence represents a mouse sonic  
 CC hedgehog (Shh) polypeptide.  
 XX

SQ Sequence 437 AA;

Query Match 100.0%; Score 2289; DB 22; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-226;  
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLILARCFVLIVLASSLVCPGLACGPGRGCKRRHKPLPLAVKQIPVNAEKTIGAS 60  
 Db 1 MLILARCFVLIVLASSLVCPGLACGPGRGCKRRHKPLPLAVKQIPVNAEKTIGAS 60  
 QY 61 GRYECKITRNSERKELTPYNPDIIFDEENTGADRLMTORCKDKLNAALISVNWMPG 120  
 Db 61 GRYECKITRNSERKELTPYNPDIIFDEENTGADRLMTORCKDKLNAALISVNWMPG 120  
 QY 121 VRLVTEGMDGDGHHSESLHYEGRAVDITTSDDRSRYGMLARLAVAGFDWYYESKA 180  
 Db 121 VRLVTEGMDGDGHHSESLHYEGRAVDITTSDDRSRYGMLARLAVAGFDWYYESKA 180

QY 181 HHCYKAENSVAKSGCGPGSATYHLEGGTKLYKDLRPGDRVLAADQGRILYSDFL 240  
 Db 181 HHCYKAENSVAKSGCGPGSATYHLEGGTKLYKDLRPGDRVLAADQGRILYSDFL 240  
 QY 241 TELDRDEGAKKYFYVETLEPRERLLTAAHLEFVAPHNDSGTPSPSLFASRPNGOR 300  
 Db 241 TELDRDEGAKKYFYVETLEPRERLLTAAHLEFVAPHNDSGTPSPSLFASRPNGOR 300  
 QY 301 VYVVAERGGDRLLPAAVSVTLREBEAGAVAPLTAHGTILLNRVLASCYAVIEERSMAH 360  
 Db 301 VYVVAERGGDRLLPAAVSVTLREBEAGAVAPLTAHGTILLNRVLASCYAVIEERSMAH 360  
 QY 361 RAFAFRLAHALLAALAPARTDGGGGSIIPAQSATEARGAEPAGIHWYSQLYHIGTW 420  
 Db 361 RAFAFRLAHALLAALAPARTDGGGGSIIPAQSATEARGAEPAGIHWYSQLYHIGTW 420  
 QY 421 LLDSETHMPLGMAVKSS 437  
 Db 421 LLDSETHMPLGMAVKSS 437

RESULT 11

AAB85736  
 ID AAB85736 standard; Protein: 437 AA.  
 AC AAB85736;  
 XX  
 DT 29-OCT-2001 (first entry)  
 XX  
 DE Mouse sonic hedgehog (Shh) polypeptide.  
 XX  
 KW Hedgehog protein; sonic hedgehog; Shh; indian hedgehog; Ihh; Dhh;  
 XX desert hedgehog; cell differentiation; mouse.  
 OS Mus sp.  
 XX  
 PN US6271363-B1.  
 XX  
 PD 07-AUG-2001.  
 XX  
 PF 20-OCT-1997; 97US-0954698.  
 XX  
 PR 05-JUN-1995; 95US-0462386.  
 XX  
 PR 30-DEC-1993; 93US-0176427.  
 XX  
 PR 14-DEC-1994; 94US-0356060.  
 XX  
 PR 04-MAY-1995; 95US-0435093.  
 XX

PA (HARD ) HARVARD COLLEGE.  
 PI (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
 XX  
 DR Ingham PW, McMahon AP, Tablin CJ;  
 XX  
 DR WPI: 2001-456723/49.  
 XX  
 DR N-PSDB; AAH76110.  
 XX

Novel nucleic acid encoding a hedgehog polypeptide, used to produce the  
 polypeptide, which is used to promote proliferation, survival, and/or  
 differentiation of neuronal and mesodermal tissue -  
 Claim 1; Column 127-130; 118pp; English.

CC The invention relates to nucleic acids encoding hedgehog proteins  
 CC selected from sonic hedgehog (Shh), indian hedgehog (Ihh), desert  
 CC hedgehog (Dhh) polypeptides. The hedgehog genes are involved in the  
 CC formation of ordered spatial arrangements of differentiated tissue in  
 CC vertebrates. The nucleic acid sequences are useful for producing hedgehog  
 CC proteins, used for promoting differentiation of, or survival of  
 CC or differentiation of mesenchymal, endodermal or ectodermal tissue,  
 CC particularly chondrocytes, or testicular germ line cells. The present  
 CC sequence represents a mouse Shh polypeptide.  
 XX

SO Sequence 437 AA: 100.0%; Score 2289; DB 22; Length 437;  
 Query Match 100.0%; Pred. No. 9.4e-226;  
 Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;  
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLLVCPGLACGPGRGGRKRRHKRLTPLAYKQFIPNVAEKTIGAS 60  
 1 MLLIARCFVLIASSLLVCPGLACGPGRGGRKRRHKRLTPLAYKQFIPNVAEKTIGAS 60  
 DB 1 MLLIARCFVLIASSLLVCPGLACGPGRGGRKRRHKRLTPLAYKQFIPNVAEKTIGAS 60  
 QY 61 GRVSGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALISVNMQWPG 120  
 61 GRVSGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALISVNMQWPG 120  
 DB 61 GRVSGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALISVNMQWPG 120  
 QY 121 VRLRTEGMDDEGHSEESLHYEGRAVDITTSDBRSKYGMRLARLAVEAGFDWYYESKA 180  
 121 VRLRTEGMDDEGHSEESLHYEGRAVDITTSDBRSKYGMRLARLAVEAGFDWYYESKA 180  
 DB 121 VRLRTEGMDDEGHSEESLHYEGRAVDITTSDBRSKYGMRLARLAVEAGFDWYYESKA 180  
 QY 181 HIRCSVKAENSVAAKSGCGPFSATVHLEOGTKLVKDLRPGDRVLADDDGRLYSDFL 240  
 181 HIRCSVKAENSVAAKSGCGPFSATVHLEOGTKLVKDLRPGDRVLADDDGRLYSDFL 240  
 DB 181 HIRCSVKAENSVAAKSGCGPFSATVHLEOGTKLVKDLRPGDRVLADDDGRLYSDFL 240  
 QY 241 TFLDRDEGAKKVFYIETLEPERERLLTPAHLLEFVAPHNDSGPTGPSALFASRVPGQR 300  
 241 TFLDRDEGAKKVFYIETLEPERERLLTPAHLLEFVAPHNDSGPTGPSALFASRVPGQR 300  
 DB 241 TFLDRDEGAKKVFYIETLEPERERLLTPAHLLEFVAPHNDSGPTGPSALFASRVPGQR 300  
 QY 301 VYVVAERGGDRRLPAVAHSVTLREEDAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360  
 301 VYVVAERGGDRRLPAVAHSVTLREEDAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360  
 DB 301 VYVVAERGGDRRLPAVAHSVTLREEDAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360  
 QY 361 RAFAFPRLLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWSQLYHIGTW 420  
 361 RAFAFPRLLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWSQLYHIGTW 420  
 DB 361 RAFAFPRLLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWSQLYHIGTW 420  
 QY 421 LLDSETHPLGMAVKSS 437  
 421 LLDSETHPLGMAVKSS 437  
 DB 421 LLDSETHPLGMAVKSS 437

RESULT 12  
 AAB84672  
 ID AAB84672 standard; Protein; 437 AA.  
 XX AAB84672;  
 AC 17-SEP-2001 (first entry)  
 XX  
 DT  
 XX  
 DE Amino acid sequence of a mouse hedgehog (Shh) polypeptide.  
 XX  
 KW Shh gene; hedgehog gene; T lymphocyte; patched gene; infection;  
 KW diabetes; nutritional deficiency; graft rejection; hyperacute response;  
 KW cornea transplant; autoimmune disorder; multiple sclerosis; psoriasis;  
 KW atopic dermatitis; inflammatory disease; proliferative disease;  
 KW hyperproliferative disease; eczematous dermatitis; urticaria;  
 KW vasculitis; scleroderma.  
 XX  
 KW  
 OS Mus sp.  
 XX  
 PN WO200140438-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 30-NOV-2000; 2000MO-US32590.  
 XX  
 PR 30-NOV-1999; 99US-0168112.  
 XX  
 PA (CUR1-) CURIS INC.  
 XX  
 PI Crompton T;  
 XX  
 DR WPI: 2001-441484/47.  
 XX  
 DR N-PSDB; AAB84672.  
 XX

PT Modulating immune function comprises administration of a hedgehog or  
 PT agent, for treating e.g. diabetes, eczematous dermatitis, urticaria  
 PT or vasculitis -  
 XX  
 PS Claim 4; Page 81-82; 105pp; English.

The present sequence represents a hedgehog (Shh) polypeptide. Hedgehog  
 CC gene products and signal transduction pathways involving hedgehog are  
 CC involved in the maturation of T lymphocytes. The specification describes  
 CC a method for modulating immune function, by administration of a hedgehog  
 CC or patched (ptc) polypeptide, agonists or antagonists. The method is  
 CC used to treat disorders affecting the regulation of lymphocytes,  
 CC particularly maturation and/or activation of T lymphocytes. It is used  
 CC to treat bacterial or viral infection, diabetes, nutritional  
 CC deficiencies, graft rejection or other hyperacute response such as  
 CC kidney, heart, lung, bone marrow spleen skin or cornea transplant or  
 CC autoimmune disorders such as multiple sclerosis, psoriasis or atopic  
 CC dermatitis. The method is used to treat inflammatory, proliferative and  
 CC hyperproliferative diseases, as well as cutaneous manifestations of  
 CC immunological disorders such as eczematous dermatitis, urticaria,  
 CC vasculitis and scleroderma.

SO Sequence 437 AA: 100.0%; Score 2289; DB 22; Length 437;  
 Query Match 100.0%; Pred. No. 9.4e-226;  
 Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;  
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLLVCPGLACGPGRGGRKRRHKRLTPLAYKQFIPNVAEKTIGAS 60  
 1 MLLIARCFVLIASSLLVCPGLACGPGRGGRKRRHKRLTPLAYKQFIPNVAEKTIGAS 60  
 DB 1 MLLIARCFVLIASSLLVCPGLACGPGRGGRKRRHKRLTPLAYKQFIPNVAEKTIGAS 60  
 QY 61 GRVSGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALISVNMQWPG 120  
 61 GRVSGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALISVNMQWPG 120  
 DB 61 GRVSGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALISVNMQWPG 120  
 QY 121 VRLRTEGMDDEGHSEESLHYEGRAVDITTSDBRSKYGMRLARLAVEAGFDWYYESKA 180  
 121 VRLRTEGMDDEGHSEESLHYEGRAVDITTSDBRSKYGMRLARLAVEAGFDWYYESKA 180  
 DB 121 VRLRTEGMDDEGHSEESLHYEGRAVDITTSDBRSKYGMRLARLAVEAGFDWYYESKA 180  
 QY 181 HIRCSVKAENSVAAKSGCGPFSATVHLEOGTKLVKDLRPGDRVLADDDGRLYSDFL 240  
 181 HIRCSVKAENSVAAKSGCGPFSATVHLEOGTKLVKDLRPGDRVLADDDGRLYSDFL 240  
 DB 181 HIRCSVKAENSVAAKSGCGPFSATVHLEOGTKLVKDLRPGDRVLADDDGRLYSDFL 240  
 QY 241 TFLDRDEGAKKVFYIETLEPERERLLTPAHLLEFVAPHNDSGPTGPSALFASRVPGQR 300  
 241 TFLDRDEGAKKVFYIETLEPERERLLTPAHLLEFVAPHNDSGPTGPSALFASRVPGQR 300  
 DB 241 TFLDRDEGAKKVFYIETLEPERERLLTPAHLLEFVAPHNDSGPTGPSALFASRVPGQR 300  
 QY 301 VYVVAERGGDRRLPAVAHSVTLREEDAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360  
 301 VYVVAERGGDRRLPAVAHSVTLREEDAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360  
 DB 301 VYVVAERGGDRRLPAVAHSVTLREEDAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360  
 QY 361 RAFAFPRLLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWSQLYHIGTW 420  
 361 RAFAFPRLLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWSQLYHIGTW 420  
 DB 361 RAFAFPRLLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWSQLYHIGTW 420  
 QY 421 LLDSETHPLGMAVKSS 437  
 421 LLDSETHPLGMAVKSS 437  
 DB 421 LLDSETHPLGMAVKSS 437

RESULT 13  
 AAE04685  
 ID AAE04685 standard; Protein; 437 AA.  
 XX AAE04685;  
 AC 04-SEP-2001 (first entry)  
 XX  
 DT  
 XX  
 DE Mouse sonic hedgehog (Shh) protein.  
 XX  
 KW Mouse; hedgehog protein; noctropic; neuroprotective; anticonvulsant;  
 KW cytoskeletal; therapy; Alzheimer's disease; Parkinson's disease; injury;  
 KW

KW Huntington's chorea; amyotrophic lateral sclerosis; multiple sclerosis;  
 KW nervous system aging; neurodegenerative disease; immunological disease;  
 KW malignant glioma; medulloblastoma; neuroectodermal tumour; cancer;  
 KW extracellular signalling protein.  
 OS Mus sp.  
 XX WO200134654-A1.  
 XX PD 17-MAY-2001.  
 XX PF 02-NOV-2000; 2000WO-US30405.  
 XX PR 05-NOV-1999; 99US-0164025.  
 XX PA (BIOJ ) BIOGEN INC.  
 XX PI Strauch K;  
 XX DR WPI: 2001-329075/34.  
 XX DR N-PSDB; AAD09032.  
 XX PT Novel isolated hedgehog fusion polypeptide useful for treating  
 PT neurological conditions such as Alzheimer's disease, Parkinson's  
 PT disease, Huntington's chorea, amyotrophic lateral sclerosis, and  
 XX multiple sclerosis.  
 PS Disclosure; Page 115-116; 178pp; English.  
 CC The present invention relates to hedgehog fusion proteins. Hedgehog  
 CC proteins are a family of extracellular signalling proteins that regulate  
 CC various aspects of embryonic development both in vertebrates and in  
 CC invertebrates. Hedgehog fusion protein is useful for the prophylaxis or  
 CC treatment of any condition or disease state for which the hedgehog or  
 CC patched protein constituent is efficacious and in the diagnosis of  
 CC constituents or conditions of disease states and in the diagnosis of  
 CC specimens and for diagnostic purposes in non-physiological systems or  
 CC Hedgehog fusion protein is useful for treating neurological conditions  
 CC due to injury, aging of nervous system, including Alzheimer's disease,  
 CC chronic neurodegenerative diseases of the nervous system, including  
 CC Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis  
 CC and chronic immunological diseases of the nervous system including multiple  
 CC sclerosis and malignant gliomas, medulloblastomas, neuroectodermal  
 CC tumours and to specifically target medical therapies against cancers and  
 CC tumours which express the receptor for the protein. The present sequence  
 CC is mouse sonic hedgehog (Shh) protein.  
 CC XX  
 SQ Sequence 437 AA:  
 Query Match 100.0%; Score 2289; DB 22; Length 437;  
 Best Local Similarity 100.0%; Pred. NO. 9.4e-226;  
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLLILARCFVLIASSLLVCPGLACGPGRGFGKRRPKKLPLAVYKQFIPVNAKTIIGAS 60  
 DB 1 MLLILARCFVLIASSLLVCPGLACGPGRGFGKRRPKKLPLAVYKQFIPVNAKTIIGAS 60  
 QY 61 GRECKITRNSERKELTPYNDPIIFKDEMGADRLMTORCKDKALAISSWQMPG 120  
 DB 61 GRECKITRNSERKELTPYNDPIIFKDEMGADRLMTORCKDKALAISSWQMPG 120  
 QY 121 VRLRVTEGMDHSEESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDMVYVESKA 180  
 DB 121 VRLRVTEGMDHSEESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDMVYVESKA 180  
 QY 181 HHCSTKKAANSVAANSKSGCGPGSATVHLFEGGTKLVKDLRPDQRYLAADQGRLLYSDFL 240  
 DB 181 HHCSTKKAANSVAANSKSGCGPGSATVHLFEGGTKLVKDLRPDQRYLAADQGRLLYSDFL 240  
 QY 241 TFLDDEGAKKVFYIETLEPERLLTAHLLEFVAPHNDSPPTGSPALFASRVPRGQR 300  
 DB 241 TFLDDEGAKKVFYIETLEPERLLTAHLLEFVAPHNDSPPTGSPALFASRVPRGQR 300

QY 301 VYVVAERGDRRLPLPAVHSVTLREBAGAYAPLTAHGTILINRVLASCYAVEESHMAH 360  
 DB 301 VYVVAERGDRRLPLPAVHSVTLREBAGAYAPLTAHGTILINRVLASCYAVEESHMAH 360  
 QY 361 RAFAFRLAHALLAALAPARTDGGGSGTIPAAQSATBANGAEPPTAGIHWSQLLYHIGTW 420  
 DB 361 RAFAFRLAHALLAALAPARTDGGGSGTIPAAQSATBANGAEPPTAGIHWSQLLYHIGTW 420  
 QY 421 LLDSETHMPLGMAVKS 437  
 DB 421 LLDSETHMPLGMAVKS 437  
 RESULT 14  
 AAE05375  
 ID AAE05375 standard; Protein: 437 AA.  
 AC AAE05375;  
 XX 12-SEP-2001 (first entry)  
 DE Mouse Sonic hedgehog protein.  
 KW Mouse; Sonic hedgehog; Shh; morphogenic signal; neuron; chromosome 5;  
 KW embryonic patterning; cell culture; cell differentiation; ischaemia;  
 KW cell proliferative disorder; intracerebral grafting; Huntington's chorea;  
 KW neurological disorder; Alzheimer's disease; Parkinson's disease;  
 KW amyotrophic lateral sclerosis; ALS; multiple sclerosis.  
 OS Mus sp.  
 XX  
 XX Key Location/Qualifiers  
 XX Peptide 1..24  
 XX Protein /label=Signal\_peptide  
 XX /label=25..437  
 XX /label=Mature\_Shh\_protein  
 XX US6261786-B1.  
 XX 17-JUL-2001.  
 XX 02-JUL-1996; 96US-0674509.  
 XX 30-DEC-1993; 93US-0176427.  
 XX 14-DEC-1994; 94US-0356060.  
 XX 04-MAY-1995; 95US-0435093.  
 XX 05-JUN-1995; 95US-0460900.  
 XX 05-JUN-1995; 95US-0462386.  
 PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
 PA (HARD ) HARVARD COLLEGE.  
 PI Mariago V, Tablin CJ, Ingham PW, McMahon AP;  
 XX WPI: 2001-440859/47.  
 XX N-PSDB; AAD10149.  
 PT Screening compounds that potentiate or inhibit binding of hedgehog  
 PT polypeptide to naturally occurring patched receptor, comprises  
 PT contacting polypeptide with receptor and test compound, and detecting  
 PT change in binding.  
 PS Claim 2; Fig 5A; 127pp; English.  
 CC The present invention relates to assay for screening compounds that  
 CC potentiate or inhibit binding of hedgehog polypeptide to naturally  
 CC occurring patched receptor. The hedgehog proteins comprise morphogenic  
 CC signals produced by embryonic patterning centres, and are involved in  
 CC formation and maintenance of ordered spatial arrangements of  
 CC differentiated tissues in vertebrates, both adult and embryonic. The  
 CC proteins can be used to generate and/or maintain an array of different  
 CC vertebrate tissues both in vitro and in vivo. The invention also relates  
 CC to a method for modulating growth, differentiation or survival of a

CC mammalian cell (e.g. neuron, testicular cell) responsive to hedgehog  
 CC induction. Hedgehog agonists and antagonists can be used in cell culture  
 CC techniques to enhance survival and maintenance of neurons and various  
 CC vertebrate organogenic pathways. The hedgehog gene is useful in  
 CC determining whether a patient is at the risk of disorder characterised by  
 CC unwanted cell proliferation or aberrant control of differentiation. The  
 CC hedgehog proteins or mimetics can be used to induce fetal neurons  
 CC especially neuronal stem cells in intracerebral grafting. The protein  
 CC or its mimetic can be used in the treatment of neurological conditions  
 CC e.g. injury to nervous system, ischaemia resulting from stroke,  
 CC Alzheimer's disease, Parkinson's disease, Huntington's chorea,  
 CC amyotrophic lateral sclerosis (ALS) and multiple sclerosis. The present  
 CC sequence is mouse Sonic hedgehog (Shh) protein. The Shh gene is  
 CC mapped to the proximal region of mouse chromosome 5.

SO Sequence 437 AA:

Query Match 100.0%: Score 2289; DB 22; Length 437;

Best Local Similarity 100.0%: Pred. No. 9.4e-226; Mismatches 0; Indels 0; Gaps 0;

Matches 437: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLARCFVLVIASSLVCPLACPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60  
 DB 1 MLLLARCFVLVIASSLVCPLACPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60  
 QY 61 GRYEGKITRNSERFKELTNPYNPDIIFKDEMTGADRLMTQRCCKDLNLAISVNMQMG 120  
 DB 61 GRYEGKITRNSERFKELTNPYNPDIIFKDEMTGADRLMTQRCCKDLNLAISVNMQMG 120  
 QY 121 VRLRTYEGWDEGHHSEELHTEGRAVDITTSDRDSKTKGMLARLAVEGFDMVYTESKA 180  
 DB 121 VRLRTYEGWDEGHHSEELHTEGRAVDITTSDRDSKTKGMLARLAVEGFDMVYTESKA 180  
 QY 181 HTHCSYKAENSYAASGGCGFPGSATYHLEOGGKTKLVKDLRPGDRLAADOGLLYSDPL 240  
 DB 181 HTHCSYKAENSYAASGGCGFPGSATYHLEOGGKTKLVKDLRPGDRLAADOGLLYSDPL 240  
 QY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDSGTPPGSALFASRVAPGQR 300  
 DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDSGTPPGSALFASRVAPGQR 300  
 QY 301 VYVVAERGGDRRLPRAAVSVTLREBEAGATYAPLTAHGITILNRYLASCYAVIEESHMAH 360  
 DB 301 VYVVAERGGDRRLPRAAVSVTLREBEAGATYAPLTAHGITILNRYLASCYAVIEESHMAH 360  
 QY 361 RAFAPFRLAHALLALAPARTDGGGGSIPAAQSATEARGAEPAGIHMYSQLYHIGTW 420  
 DB 361 RAFAPFRLAHALLALAPARTDGGGGSIPAAQSATEARGAEPAGIHMYSQLYHIGTW 420  
 QY 421 LLDSETHMPLGMAVKSS 437  
 DB 421 LLDSETHMPLGMAVKSS 437

RESULT 15

AAB31220 AAB31220 standard; Protein; 437 AA.

AC AAB31220;

DT 20-APR-2001 (first entry)

DE Amino acid sequence of mouse sonic hedgehog protein (Shh).

KM Hedgehog related-protein; sonic hedgehog protein; Shh; Ischemia; stroke;  
 KM desert hedgehog protein; Dhh; Indian hedgehog protein; Ihh; neuron;  
 KM neurological condition; nervous system injury; tumour-induced injury;  
 KM aging; Alzheimer's disease; chronic neurodegenerative disease;  
 KM Parkinson's disease; Huntington's chorea; amyotrophic lateral sclerosis;  
 KM spinocerebellar degeneration; chronic immunological disease;  
 KM multiple sclerosis.

OS Mus sp.

XX US6165747-A.  
 PN 26-DEC-2000.  
 PD 05-JUN-1995; 95US-0460900.  
 PF 30-DEC-1993; 93US-0176427.  
 PR 14-DEC-1994; 94US-0356060.  
 PR 04-MAY-1995; 95US-0435093.  
 PA (HARD ) HARVARD COLLEGE.  
 PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
 XX Ingham PW, McMahon AP, Tablin CJ, Marti-gorostiza E, Bumcrot DA;  
 PI WPI: 2001-079847/09.  
 DR N-PSDB; AAC87077.  
 XX Polynucleotides encoding hedgehog proteins, useful for treating  
 PT diseases of nervous system such as Alzheimer's disease, Parkinson's  
 PT disease, Huntington's chorea, amyotrophic lateral sclerosis, multiple  
 PT sclerosis -  
 PS Claim 10; Columns 143-146; 119pp; English.

The present sequence represents a hedgehog related-protein. The  
 CC specification describes a sonic hedgehog protein (Shh), a desert  
 CC hedgehog protein (Dhh), and an Indian hedgehog protein (Ihh). The  
 CC hedgehog polynucleotides are useful in diagnostic, in antisense  
 CC therapy and in therapeutic assays for detecting and treating disorders  
 CC involving, e.g., aberrant expression of vertebrate hedgehog homologue.  
 CC Hedgehog polypeptides are useful therapeutically to enhance survival  
 CC of neurons and other neuron cells and in treating neurological  
 CC conditions deriving from acute, subacute, or chronic injury to the  
 CC nervous system, including traumatic injury, chemical injury, vascular  
 CC injury and deficits (such as the ischemia resulting from stroke),  
 CC together with infectious/inflammatory and induced-induced injury, aging  
 CC of the nervous system including Alzheimer's disease, chronic  
 CC neurodegenerative diseases of the nervous system, including Parkinson's  
 CC disease, Huntington's chorea, amyotrophic lateral sclerosis,  
 CC spinocerebellar degenerations, and chronic immunological diseases of  
 CC the nervous system or affecting the nervous system, including multiple  
 CC sclerosis.

SO Sequence 437 AA:

Query Match 100.0%: Score 2289; DB 22; Length 437;

Best Local Similarity 100.0%: Pred. No. 9.4e-226; Mismatches 0; Indels 0; Gaps 0;

Matches 437: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLARCFVLVIASSLVCPLACPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60  
 DB 1 MLLLARCFVLVIASSLVCPLACPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60  
 QY 61 GRYEGKITRNSERFKELTNPYNPDIIFKDEMTGADRLMTQRCCKDLNLAISVNMQMG 120  
 DB 61 GRYEGKITRNSERFKELTNPYNPDIIFKDEMTGADRLMTQRCCKDLNLAISVNMQMG 120  
 QY 121 VRLRTYEGWDEGHHSEELHTEGRAVDITTSDRDSKTKGMLARLAVEGFDMVYTESKA 180  
 DB 121 VRLRTYEGWDEGHHSEELHTEGRAVDITTSDRDSKTKGMLARLAVEGFDMVYTESKA 180  
 QY 181 HTHCSYKAENSYAASGGCGFPGSATYHLEOGGKTKLVKDLRPGDRLAADOGLLYSDPL 240  
 DB 181 HTHCSYKAENSYAASGGCGFPGSATYHLEOGGKTKLVKDLRPGDRLAADOGLLYSDPL 240  
 QY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDSGTPPGSALFASRVAPGQR 300  
 DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDSGTPPGSALFASRVAPGQR 300  
 QY 301 VYVVAERGGDRRLPRAAVSVTLREBEAGATYAPLTAHGITILNRYLASCYAVIEESHMAH 360  
 DB 301 VYVVAERGGDRRLPRAAVSVTLREBEAGATYAPLTAHGITILNRYLASCYAVIEESHMAH 360

Db 301 VYVAERCGDRRLPAVAHSTVLRREEAGAYAPLTAHCTILINRVIASCYAVIEHSMW 360  
Oy 361 RAFAFERLAHALALAPARTDGGGGSIPMAOSATEARGAEPYAGIHWSQLYHGTW 420  
Db 361 RAFAFERLAHALALAPARTDGGGGSIPMAOSATEARGAEPYAGIHWSQLYHGTW 420  
Oy 421 LLDSETHMPLGMAVKSS 437  
Db 421 LLDSETHMPLGMAVKSS 437

Search completed: February 20, 2003, 10:10:18  
Job time : 32.7046 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:08:38 ; Search time 11.4398 Seconds  
(without alignments) 1123.956 Million cell updates/sec

Title: US-09-827-110A-13  
Perfect score: 2289  
Sequence: 1 MLLLRARCFVLIASLLVC.....CTWLDSFTMPLGMAYKSS 437

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues  
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2289	100.0	437	1 US-08-176-427B-8	Sequence 8, Appl1
2	2289	100.0	437	2 US-08-356-060A-11	Sequence 11, Appl1
3	2289	100.0	437	4 US-08-460-900C-11	Sequence 11, Appl1
4	2289	100.0	437	4 US-08-674-509B-11	Sequence 11, Appl1
5	2289	100.0	437	4 US-08-954-698-11	Sequence 11, Appl1
6	2289	100.0	437	4 US-08-957-874-11	Sequence 11, Appl1
7	2289	100.0	437	4 US-09-325-256-20	Sequence 20, Appl1
8	2286	99.9	437	3 US-08-946-329A-20	Sequence 20, Appl1
9	2286	99.9	437	4 US-08-567-357A-20	Sequence 20, Appl1
10	2286	99.9	437	4 US-08-729-743A-20	Sequence 20, Appl1
11	2286	99.9	437	4 US-08-057-860A-6	Sequence 6, Appl1
12	2286	99.9	437	4 US-08-349-498-20	Sequence 20, Appl1
13	2286	99.9	437	4 US-08-293-505-14	Sequence 14, Appl1
14	2286	99.9	437	5 PCT-US95-15463-20	Sequence 20, Appl1
15	2286	99.9	437	5 PCT-US95-15923-20	Sequence 20, Appl1
16	2200	96.1	437	4 US-08-757-230A-2	Sequence 2, Appl1
17	2200	96.1	437	4 US-08-757-230A-9	Sequence 9, Appl1
18	2200	96.1	437	5 PCT-US95-02315-2	Sequence 2, Appl1
19	2065	90.2	462	1 US-08-748-591-4	Sequence 9, Appl1
20	2065	90.2	462	1 US-08-356-060A-13	Sequence 13, Appl1
21	2065	90.2	475	2 US-08-460-900C-13	Sequence 13, Appl1
22	2065	90.2	475	4 US-08-674-509B-13	Sequence 13, Appl1
23	2065	90.2	475	4 US-08-954-698-13	Sequence 13, Appl1
24	2065	90.2	475	4 US-08-957-874-13	Sequence 13, Appl1
25	2065	90.2	475	4 US-09-325-256-22	Sequence 22, Appl1
26	2065	90.2	475	4 US-08-349-498-18	Sequence 18, Appl1
27	1788	78.1	425	1 US-08-176-427B-2	Sequence 2, Appl1

28	1788	78.1	425	2 US-08-356-060A-8	Sequence 8, Appl1
29	1788	78.1	425	4 US-08-460-900C-8	Sequence 8, Appl1
30	1788	78.1	425	4 US-08-674-509B-8	Sequence 8, Appl1
31	1788	78.1	425	4 US-08-954-698-8	Sequence 8, Appl1
32	1788	78.1	425	4 US-08-957-874-8	Sequence 8, Appl1
33	1788	78.1	425	4 US-09-325-256-17	Sequence 17, Appl1
34	1679	73.4	425	3 US-08-567-357A-19	Sequence 19, Appl1
35	1679	73.4	425	4 US-08-729-743A-19	Sequence 19, Appl1
36	1679	73.4	425	4 US-08-349-498-19	Sequence 19, Appl1
37	1679	73.4	425	4 US-08-15463-19	Sequence 19, Appl1
38	1679	73.4	425	5 PCT-US95-15923-19	Sequence 19, Appl1
39	1679	73.4	425	5 US-08-946-329A-18	Sequence 18, Appl1
40	1517	66.3	418	3 US-08-567-357A-18	Sequence 18, Appl1
41	1517	66.3	418	4 US-08-729-743A-18	Sequence 18, Appl1
42	1517	66.3	418	4 US-08-349-498-18	Sequence 18, Appl1
43	1517	66.3	418	4 US-08-757-230A-7	Sequence 7, Appl1
44	1517	66.3	418	4 US-08-349-498-18	Sequence 18, Appl1
45	1517	66.3	418	5 PCT-US95-15463-18	Sequence 18, Appl1

## ALIGNMENTS

```

RESULT 1
US-08-176-427B-8
; Sequence 8, Application US/08176427B
; Patent No. 5789543
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHYE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-427B-8
Query Match      100.0%  Score 2289; DB 1; Length 437;
Best local Similarity 100.0%; Pred. No. 2.7e-249; Indels 0; Gaps 0;
Matches 437; Conservative 0; Mismatches 0;
QY 1 MLLLRARCFVLIASLLVCPGLACPGRGFGRRHKLTPLAYKQFIIPNVAEKTGLAS 60
DB 1 MLLLRARCFVLIASLLVCPGLACPGRGFGRRHKLTPLAYKQFIIPNVAEKTGLAS 60
  
```

QY 61 GRYECKITRNSERFELTPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQMG 120  
DB 61 GRYECKITRNSERFELTPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQMG 120  
QY 121 VRLRTEGDEDEGHSESLHYEGRAVDITTSDBRDSKYGMLARLAVEAGFDWYYESKA 180  
DB 121 VRLRTEGDEDEGHSESLHYEGRAVDITTSDBRDSKYGMLARLAVEAGFDWYYESKA 180  
QY 181 HHC SVKAENSVAAKSGGCGPSATVHLEOGGKLVKDLRPGDRVLAADDOGRILYSDFL 240  
DB 181 HHC SVKAENSVAAKSGGCGPSATVHLEOGGKLVKDLRPGDRVLAADDOGRILYSDFL 240  
QY 241 TFLDRDGAKKVFYVETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300  
DB 241 TFLDRDGAKKVFYVETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300  
QY 301 VYVAERGGDRLLPAAVHSVTLREEAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360  
DB 301 VYVAERGGDRLLPAAVHSVTLREEAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360  
QY 361 RAFAFRLLHALALAPARTDGGGGSIIPAOSATEARGAEPGTAGIMYSOLLHYIGTW 420  
DB 361 RAFAFRLLHALALAPARTDGGGGSIIPAOSATEARGAEPGTAGIMYSOLLHYIGTW 420  
QY 421 LLDSETHMPLGMAVKSS 437  
DB 421 LLDSETHMPLGMAVKSS 437

RESULT 2  
US-08-356-060A-11  
Sequence 11, Application US/08356060A  
Patent No. 5844079  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tablin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356,060A  
FILING DATE: 14-DEC-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMT-006CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-356-060A-11

Query Match 100.0%; Score 2289; DB 2; Length 437;  
Best Local Similarity 100.0%; Pred. No. 2, 7e-249;  
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLLIARCFVLVIASSLIVCPGLACGPRGGRKRLPLAYKOFINVAEKTIGAS 60  
DB 1 MLLIARCFVLVIASSLIVCPGLACGPRGGRKRLPLAYKOFINVAEKTIGAS 60  
QY 61 GRYECKITRNSERFELTPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQMG 120  
DB 61 GRYECKITRNSERFELTPNYNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQMG 120  
QY 121 VRLRTEGDEDEGHSESLHYEGRAVDITTSDBRDSKYGMLARLAVEAGFDWYYESKA 180  
DB 121 VRLRTEGDEDEGHSESLHYEGRAVDITTSDBRDSKYGMLARLAVEAGFDWYYESKA 180  
QY 181 HHC SVKAENSVAAKSGGCGPSATVHLEOGGKLVKDLRPGDRVLAADDOGRILYSDFL 240  
DB 181 HHC SVKAENSVAAKSGGCGPSATVHLEOGGKLVKDLRPGDRVLAADDOGRILYSDFL 240  
QY 241 TFLDRDGAKKVFYVETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300  
DB 241 TFLDRDGAKKVFYVETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300  
QY 301 VYVAERGGDRLLPAAVHSVTLREEAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360  
DB 301 VYVAERGGDRLLPAAVHSVTLREEAGAYAPLTAHGTLINRVLASCYAVIEESHMAH 360  
QY 361 RAFAFRLLHALALAPARTDGGGGSIIPAOSATEARGAEPGTAGIMYSOLLHYIGTW 420  
DB 361 RAFAFRLLHALALAPARTDGGGGSIIPAOSATEARGAEPGTAGIMYSOLLHYIGTW 420  
QY 421 LLDSETHMPLGMAVKSS 437  
DB 421 LLDSETHMPLGMAVKSS 437

RESULT 3  
US-08-460-900C-11  
Sequence 11, Application US/08460900C  
Patent No. 6165747  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tablin, Clifford J.  
APPLICANT: Bumcrot, David A.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,900C  
FILING DATE: 5-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 4-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
PRIOR APPLICATION DATA:

```

: APPLICATION NUMBER: US 08/176,427
: FILING DATE: 30-DEC-1993
: ATTORNEY/AGENT INFORMATION:
:   NAME: Vincent, Matthew P.
:   REGISTRATION NUMBER: 36,709
:   REFERENCE/DOCKET NUMBER: HMV-006.05
:   TELECOMMUNICATION INFORMATION:
:     TELEPHONE: (617) 832-1000
:     TELEFAX: (617) 832-7000
:   INFORMATION FOR SEQ ID NO: 11:
:     SEQUENCE CHARACTERISTICS:
:       LENGTH: 437 amino acids
:       TYPE: amino acid
:       TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-460-900C-11

Query Match      100.0%; Score 2289; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.7e-249;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLARCFVLVLASSLLVPCGLACPGCGFGKRRHKKPLTPLAYKQFIPIVAEKTIGAS 60
Db 1 MLLLARCFVLVLASSLLVPCGLACPGCGFGKRRHKKPLTPLAYKQFIPIVAEKTIGAS 60
QY 61 GRYEKKTRNSEREFELTPNTNPDIIFKDEENTGADRLMTORCKDKLNLAISVNMWFG 120
Db 61 GRYEKKTRNSEREFELTPNTNPDIIFKDEENTGADRLMTORCKDKLNLAISVNMWFG 120
QY 121 VRLRTGEGDEGHHSESLHYEGSAVDITTSDRDRSKYGMALRLAVEGDFPMYYESKA 180
Db 121 VRLRTGEGDEGHHSESLHYEGSAVDITTSDRDRSKYGMALRLAVEGDFPMYYESKA 180
QY 181 HIHCYKAENSVAAKSGCGPFSATVHLQGGTKLVKDLRPQDRVLAADDQRLLYSDPL 240
Db 181 HIHCYKAENSVAAKSGCGPFSATVHLQGGTKLVKDLRPQDRVLAADDQRLLYSDPL 240
QY 241 TFLDDECAKKVYIETLEPRERILLTAHILLFVAAPHNDSPPTGPSALFASRVRRPGOR 300
Db 241 TFLDDECAKKVYIETLEPRERILLTAHILLFVAAPHNDSPPTGPSALFASRVRRPGOR 300
QY 301 VYVVAERGGDRRLTPAAVHSVTLREEEGAVAPLTAHGFTLLNRYLASCYAVIEEHSMAH 360
Db 301 VYVVAERGGDRRLTPAAVHSVTLREEEGAVAPLTAHGFTLLNRYLASCYAVIEEHSMAH 360
QY 361 RAFAFRILAHALLALAPARTDGGGGSIPAAQSATEARGAEPRTAGIHMYSQLYHIGTW 420
Db 361 RAFAFRILAHALLALAPARTDGGGGSIPAAQSATEARGAEPRTAGIHMYSQLYHIGTW 420
QY 421 LIDSETMHPILGMVAKSS 437
Db 421 LIDSETMHPILGMVAKSS 437

RESULT 4
US-08-674-509B-11
: Sequence 11, Application US/08674509B
: Patent No. 6261786
: GENERAL INFORMATION:
:   APPLICANT: Ingham, Phillip W.
:   APPLICANT: McManon, Andrew P.
:   APPLICANT: Tablin, Clifford J.
:   APPLICANT: Marigo, Valeria
:   TITLE OF INVENTION: SCREENING ASSAYS FOR HEDGSGOG AGONISTS
:   TITLE OF INVENTION: AND ANTAAGONISTS
:   NUMBER OF SEQUENCES: 48
:   CORRESPONDENCE ADDRESS:
:     ADDRESS: FOLEY, HOAG & ELIOT LLP
:     STREET: One Post Office Square
:     CITY: Boston
:     STATE: MA
:     COUNTRY: USA
:     ZIP: 02109-2170

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/674,509B
FILING DATE: 02-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,900
FILING DATE: 05-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HW-006,06
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-674-509B-11

Query Match          100.0%; Score 2289; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 2,7e-249;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1  MLIIARCFVLIVLASSLLVPCGLACGPGRGFGKRRHPKRLPLAYKQFIPIVAERTLGAS 60
DB  1  MLIIARCFVLIVLASSLLVPCGLACGPGRGFGKRRHPKRLPLAYKQFIPIVAERTLGAS 60
QY  61  GRYEKKIRNRSERREKELTPNNYNDIIFDEENTGADRLMTQCRKLNALISVNWQPG 120
DB  61  GRYEKKIRNRSERREKELTPNNYNDIIFDEENTGADRLMTQCRKLNALISVNWQPG 120
QY  121  VLRARTEGMDDEGDHSEESLIHYEGHAYDITTSRDRSKYGMARLAVEAGFPMYYESKA 180
DB  121  VLRARTEGMDDEGDHSEESLIHYEGHAYDITTSRDRSKYGMARLAVEAGFPMYYESKA 180
QY  181  HIHCVKAENSVAAKSGCGFPGSATVHLEQGCTKLVKDLRPGRDVLAAADQGRLLYSDFL 240
DB  181  HIHCVKAENSVAAKSGCGFPGSATVHLEQGCTKLVKDLRPGRDVLAAADQGRLLYSDFL 240
QY  241  TFLDDEAKKVFYIETLEPRERILLTAAHLLFPAHPNDSQPTGSPSLFASRRAPGQR 300
DB  241  TFLDDEAKKVFYIETLEPRERILLTAAHLLFPAHPNDSQPTGSPSLFASRRAPGQR 300
QY  301  VYVVAERGGDRRLPAAVHSVTLREEAGAVAPLTAHGTLINRVLASCYAVIEEHSNAH 360
DB  301  VYVVAERGGDRRLPAAVHSVTLREEAGAVAPLTAHGTLINRVLASCYAVIEEHSNAH 360
QY  361  RAFAFRILAHALLAALAPARTDGGGGGSIIPAQSALEARGAEPAGIHMYSQLLVHIGTW 420
DB  361  RAFAFRILAHALLAALAPARTDGGGGGSIIPAQSALEARGAEPAGIHMYSQLLVHIGTW 420
QY  421  LLDSETMPLGMAYVSS 437
DB  421  LLDSETMPLGMAYVSS 437

RESULT 5
US-08-954-698-11
: Sequence 11, Application US/08954698
: Patent No. 6271363
: GENERAL INFORMATION:
: APPLICANT: Ingham, Phillip W.
: APPLICANT: McMahon, Andrew P.
: APPLICANT: Tablin, Clifford J.
: TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

```

```

? TITLE OF INVENTION: Proteins and Uses Related Thereto
? NUMBER OF SEQUENCES: 48
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: FOLEY, HOAG & ELIOT LLP
? STREET: One Post Office Square
? CITY: Boston
? STATE: MA
? COUNTRY: USA
? ZIP: 02109-2170
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/954,698
? FILING DATE: 20-OCT-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/462,386
? FILING DATE: 05-JUN-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/435,093
? FILING DATE: 04-MAY-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/356,060
? FILING DATE: 14-DEC-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/176,427
? FILING DATE: 30-DEC-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Vincent, Matthew P.
? REGISTRATION NUMBER: 36,709
? REFERENCE/DOCKET NUMBER: HMV-006.10
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-832-1000
? TELEFAX: 617-832-7000
? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 437 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-954-698-11

Query Match
Best Local Similarity 100.0%; Score 2289; DB 4; Length 437;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLARCFVLVIASSLVCPCGACGPGRGFGRRRPHKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLARCFVLVIASSLVCPCGACGPGRGFGRRRPHKLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GRYEKGITRNSRFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQNG 120
DB 61 GRYEKGITRNSRFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQNG 120
QY 121 VRLRTGEMDEGHHSESLHYEGNAVDITTSDDRSKYKGLARLAVAGAGDWMYEEKA 180
DB 121 VRLRTGEMDEGHHSESLHYEGNAVDITTSDDRSKYKGLARLAVAGAGDWMYEEKA 180
QY 181 HIRCSVAENSVAAKSGCGPGSATVHLEOGGKLVKDLRPGDVLAAADGGRLLYSDFL 240
DB 181 HIRCSVAENSVAAKSGCGPGSATVHLEOGGKLVKDLRPGDVLAAADGGRLLYSDFL 240
QY 241 TFLDRDGAARKVYVITTEPRERLLTAAHLFFVAPHNDSGPPGSAFASVVRPGQR 300
DB 241 TFLDRDGAARKVYVITTEPRERLLTAAHLFFVAPHNDSGPPGSAFASVVRPGQR 300
QY 301 VYVAARSGDRRLPLPAVHYVTREEAGAYAPLTAHGTLLINRVLASCAVAIEEESHMAH 360
DB 301 VYVAARSGDRRLPLPAVHYVTREEAGAYAPLTAHGTLLINRVLASCAVAIEEESHMAH 360
QY 361 RAFAFPRLAHALALAPARTDGGCGGSIIPAQASATFARGAEPYAGIHWYSOLLYHIGTW 420

```

```

DB 361 RAFAFPRLAHALALAPARTDGGCGGSIIPAQASATFARGAEPYAGIHWYSOLLYHIGTW 420
QY 421 LIDSETHPLGMVAKSS 437
DB 421 LIDSETHPLGMVAKSS 437

RESULT 6
US-08-957-874-11
? Sequence 11, Application US/08957874
? Patent No. 6384192
? GENERAL INFORMATION:
? APPLICANT: Ingham, Phillip W.
? APPLICANT: McMahon, Andrew P.
? APPLICANT: Tablin, Clifford J.
? TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
? NUMBER OF SEQUENCES: 47
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: FOLEY, HOAG & ELIOT LLP
? STREET: One Post Office Square
? CITY: Boston
? STATE: MA
? COUNTRY: USA
? ZIP: 02109
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: ASCII(text)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/957,874
? FILING DATE: 20-OCT-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/462,386
? FILING DATE: 5-JUNE-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/435,093
? FILING DATE: 4-MAY-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/356,060
? FILING DATE: 14-DEC-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/176,427
? FILING DATE: 30-DEC-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Vincent, Matthew P.
? REGISTRATION NUMBER: 36,709
? REFERENCE/DOCKET NUMBER: HMV-006.09
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 832-1000
? TELEFAX: (617) 832-7000
? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 437 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-957-874-11

Query Match
Best Local Similarity 100.0%; Score 2289; DB 4; Length 437;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLARCFVLVIASSLVCPCGACGPGRGFGRRRPHKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLARCFVLVIASSLVCPCGACGPGRGFGRRRPHKLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GRYEKGITRNSRFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQNG 120
DB 61 GRYEKGITRNSRFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALAI SVNMQNG 120

```

QY 121 VRLRVTEGDEDEGHSEESLHTEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYVESKA 180  
| | | | |  
DB 121 VRLRVTEGDEDEGHSEESLHTEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYVESKA 180  
QY 181 HHCSTVAENSVAKSGGCGPGSATVHLEGGGKLVKDLRPGDRVLAADOGRLYSDFL 240  
| | | | |  
DB 181 HHCSTVAENSVAKSGGCGPGSATVHLEGGGKLVKDLRPGDRVLAADOGRLYSDFL 240  
QY 241 TFLDRREGAKKVVYETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRVPRGQR 300  
| | | | |  
DB 241 TFLDRREGAKKVVYETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRVPRGQR 300  
QY 301 VYVAERGGDRRLPAVHSVTLREEBAGAYAPLTAHGTLLINRVLASCYAVIEESHMAH 360  
| | | | |  
DB 301 VYVAERGGDRRLPAVHSVTLREEBAGAYAPLTAHGTLLINRVLASCYAVIEESHMAH 360  
QY 361 RAFAPEFLAHLAALAAPATDGGGSIIPAOSATAREAGAEPTAGIHWYSOLLHYHIGTW 420  
| | | | |  
DB 361 RAFAPEFLAHLAALAAPATDGGGSIIPAOSATAREAGAEPTAGIHWYSOLLHYHIGTW 420  
QY 421 LDSETMHPGLMAVKSS 437  
| | | | |  
DB 421 LDSETMHPGLMAVKSS 437

RESULT 7  
US-09-325-256-20  
; Sequence 20, Application US/09325256  
; Patent No. 6444793  
; GENERAL INFORMATION:  
; APPLICANT: PEPINSKY, R. BLAKE  
; APPLICANT: BAKER, DARREN P.  
; APPLICANT: MEN, DINGYI  
; APPLICANT: WILLIAMS, KEVIN P.  
; APPLICANT: GARGER, ELLEN A.  
; APPLICANT: TAYLOR, FREDERICK R.  
; APPLICANT: GALDES, ALPHONSE  
; APPLICANT: PORTER, JEFFREY  
; TITLE OF INVENTION: HYDROPHOBICALLY-MODIFIED PROTEIN COMPOSITIONS AND  
; TITLE OF INVENTION: METHODS  
; FILE REFERENCE: BIV-067.01  
; CURRENT APPLICATION NUMBER: US/09/325,256  
; CURRENT FILING DATE: 1999-06-03  
; PRIOR APPLICATION NUMBER: 60/099,800  
; PRIOR FILING DATE: 1998-09-10  
; PRIOR APPLICATION NUMBER: 60/078,935  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/089,685  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/067,423  
; PRIOR FILING DATE: 1997-12-03  
; PRIOR APPLICATION NUMBER: PCT/US98/25676  
; PRIOR FILING DATE: 1998-12-03  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Murine sp.  
US-09-325-256-20

Query Match 100.0%; Score 2289; DB 4; Length 437;  
Best Local Similarity 100.0%; Pred. No. 2,7e-249;  
Matches 437; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 MLLIACFLVYITLASSLLVCPGLACGPRGGRGRRHKKLLPLAYKQFIPVNAKTIIGAS 60  
| | | | |  
DB 1 MLLIACFLVYITLASSLLVCPGLACGPRGGRGRRHKKLLPLAYKQFIPVNAKTIIGAS 60  
QY 61 GRVGGKTRRSEKPELTPNVNPDIIFKDEENGTGADRLMORCKDKNALAISVMNOMP 120  
| | | | |  
DB 61 GRVGGKTRRSEKPELTPNVNPDIIFKDEENGTGADRLMORCKDKNALAISVMNOMP 120

QY 121 VRLRVTEGDEDEGHSEESLHTEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYVESKA 180  
| | | | |  
DB 121 VRLRVTEGDEDEGHSEESLHTEGRAVDITTSRDRSKYGMALARLAVEAGFDWVYVESKA 180  
QY 181 HHCSTVAENSVAKSGGCGPGSATVHLEGGGKLVKDLRPGDRVLAADOGRLYSDFL 240  
| | | | |  
DB 181 HHCSTVAENSVAKSGGCGPGSATVHLEGGGKLVKDLRPGDRVLAADOGRLYSDFL 240  
QY 241 TFLDRREGAKKVVYETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRVPRGQR 300  
| | | | |  
DB 241 TFLDRREGAKKVVYETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRVPRGQR 300  
QY 301 VYVAERGGDRRLPAVHSVTLREEBAGAYAPLTAHGTLLINRVLASCYAVIEESHMAH 360  
| | | | |  
DB 301 VYVAERGGDRRLPAVHSVTLREEBAGAYAPLTAHGTLLINRVLASCYAVIEESHMAH 360  
QY 361 RAFAPEFLAHLAALAAPATDGGGSIIPAOSATAREAGAEPTAGIHWYSOLLHYHIGTW 420  
| | | | |  
DB 361 RAFAPEFLAHLAALAAPATDGGGSIIPAOSATAREAGAEPTAGIHWYSOLLHYHIGTW 420  
QY 421 LDSETMHPGLMAVKSS 437  
| | | | |  
DB 421 LDSETMHPGLMAVKSS 437

RESULT 8  
US-08-946-329A-20  
; Sequence 20, Application US/08946329A  
; Patent No. 6057091  
; GENERAL INFORMATION:  
; APPLICANT: Beachy, Philip A.  
; APPLICANT: Porter, Jeffrey A.  
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
; NUMBER OF SEQUENCES: 109  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/946,329A  
; FILING DATE: 07-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/061,323  
; FILING DATE: 07-OCT-1996  
; APPLICATION NUMBER: 08/729,743  
; FILING DATE: 10-JUL-1996  
; APPLICATION NUMBER: 08/567,357  
; FILING DATE: 04-DEC-1995  
; APPLICATION NUMBER: 08/349,498  
; FILING DATE: 02-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hallie, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/140001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 437 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-946-329A-20

Query Match 99.9%; Score 2286; DB 3; Length 437;  
 Best Local Similarity 99.8%; Pred. No. 5.9e-249;  
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

1 MLLLLARCFVLIASSLVCPGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
1 MLLLLARCFVLIASSLVCPGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
1 MLLLLARCFVLIASSLVCPGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
61 GRYEGKTRNSERKELTPVNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQMPG 120
61 GRYEGKTRNSERKELTPVNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQMPG 120
121 VRLRVTEGMDEGDHSEESLHYEGRAVDITTSDRSKYGMARLAVEAGFDWYTESKA 180
121 VRLRVTEGMDEGDHSEESLHYEGRAVDITTSDRSKYGMARLAVEAGFDWYTESKA 180
181 HHCYSKAEKNSVAAKSGGCPGSATVHLEOGGKLVKDLRPDDRVLADDOGRLLYSDFL 240
181 HHCYSKAEKNSVAAKSGGCPGSATVHLEOGGKLVKDLRPDDRVLADDOGRLLYSDFL 240
241 TFLDRREGAKKVFYIETLEPRERLLTAAHLFVAPHNDSGPTPGSALFASRVPRGQR 300
241 TFLDRREGAKKVFYIETLEPRERLLTAAHLFVAPHNDSGPTPGSALFASRVPRGQR 300
301 VYVVAERGGDRLLPAVHSTVTLREEGAGAYAPLTAGTILINRVLASCYAVIEEHSMAH 360
301 VYVVAERGGDRLLPAVHSTVTLREEGAGAYAPLTAGTILINRVLASCYAVIEEHSMAH 360
361 RAFAFRILAHALLAALAPARTDGGGSIIPAOSATEARGAEPAGIHWYSOLLHYHIGTW 420
361 RAFAFRILAHALLAALAPARTDGGGSIIPAOSATEARGAEPAGIHWYSOLLHYHIGTW 420
421 LLDSETHMPLGMAVKSS 437
421 LLDSETHMPLGMAVKSS 437

```

RESULT 9  
 US-08-567-357A-20

Sequence 20, Application US/08567357A  
 Patent No. 6132728

GENERAL INFORMATION:

APPLICANT: Beachy, Phillip A.

APPLICANT: Moon, Randall T.

TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES

NUMBER OF SEQUENCES: 37

ADDRESS: Fish & Richardson P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: FASTSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/567,357A

FILING DATE: 04-DEC-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/349,498

FILING DATE: 02-DEC-1994

ATTORNEY/AGENT INFORMATION:

NAME: Halle, Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07265/080001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 437 amino acids  
 TYPE: amino acids  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-567-357A-20

Query Match 99.9%; Score 2286; DB 4; Length 437;  
 Best Local Similarity 99.8%; Pred. No. 5.9e-249;  
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

1 MLLLLARCFVLIASSLVCPGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
1 MLLLLARCFVLIASSLVCPGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
1 MLLLLARCFVLIASSLVCPGLACGPGRGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
61 GRYEGKTRNSERKELTPVNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQMPG 120
61 GRYEGKTRNSERKELTPVNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNQMPG 120
121 VRLRVTEGMDEGDHSEESLHYEGRAVDITTSDRSKYGMARLAVEAGFDWYTESKA 180
121 VRLRVTEGMDEGDHSEESLHYEGRAVDITTSDRSKYGMARLAVEAGFDWYTESKA 180
181 HHCYSKAEKNSVAAKSGGCPGSATVHLEOGGKLVKDLRPDDRVLADDOGRLLYSDFL 240
181 HHCYSKAEKNSVAAKSGGCPGSATVHLEOGGKLVKDLRPDDRVLADDOGRLLYSDFL 240
241 TFLDRREGAKKVFYIETLEPRERLLTAAHLFVAPHNDSGPTPGSALFASRVPRGQR 300
241 TFLDRREGAKKVFYIETLEPRERLLTAAHLFVAPHNDSGPTPGSALFASRVPRGQR 300
301 VYVVAERGGDRLLPAVHSTVTLREEGAGAYAPLTAGTILINRVLASCYAVIEEHSMAH 360
301 VYVVAERGGDRLLPAVHSTVTLREEGAGAYAPLTAGTILINRVLASCYAVIEEHSMAH 360
361 RAFAFRILAHALLAALAPARTDGGGSIIPAOSATEARGAEPAGIHWYSOLLHYHIGTW 420
361 RAFAFRILAHALLAALAPARTDGGGSIIPAOSATEARGAEPAGIHWYSOLLHYHIGTW 420
421 LLDSETHMPLGMAVKSS 437
421 LLDSETHMPLGMAVKSS 437

```

RESULT 10  
 US-08-729-743A-20

Sequence 20, Application US/08729743A  
 Patent No. 6214794

GENERAL INFORMATION:

APPLICANT: Beachy, Phillip A.

APPLICANT: Moon, Randall T.

TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES

NUMBER OF SEQUENCES: 37

ADDRESS: Fish & Richardson P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: FASTSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/729,743A

FILING DATE: 07-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/567,357

FILING DATE: 04-DEC-1995

APPLICATION NUMBER: 08/349,498

FILING DATE: 02-DEC-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Halle, Lisa A.  
 REGISTRATION NUMBER: 38,347  
 REFERENCE/DOCKET NUMBER: 07265/099001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 619/678-5070  
 TELEFAX: 619/678-5099  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 437 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-729-743A-20

Query Match 99.9%; Score 2286; DB 4; Length 437;  
 Best Local Similarity 99.8%; Pred. No. 5.9e-249;  
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLARCFVLIVLASSLVCPGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTIGAS 60  
 DB 1 MLLLARCFVLIVLASSLVCPGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTIGAS 60  
 QY 61 GRYEKGITRNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNLAISVNMWPG 120  
 DB 61 GRYEKGITRNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNLAISVNMWPG 120  
 QY 121 VRLRTGEGDEGHHSESLHTEGRAVDITTSDRDSKYGMLARLAVEGFDMVYTESKA 180  
 DB 121 VRLRTGEGDEGHHSESLHTEGRAVDITTSDRDSKYGMLARLAVEGFDMVYTESKA 180  
 QY 181 HIHCVKAENSVAAKSGCGFPGSATVHLEQGTKLVDLPDGRVLAADDOGRLLYSDFL 240  
 DB 181 HIHCVKAENSVAAKSGCGFPGSATVHLEQGTKLVDLPDGRVLAADDOGRLLYSDFL 240  
 QY 241 TELDRDEGAKKVFYETLEPERELLTPAHLLEFVAPHNDSGPTGPSALFASRVAPGQR 300  
 DB 241 TELDRDEGAKKVFYETLEPERELLTPAHLLEFVAPHNDSGPTGPSALFASRVAPGQR 300  
 QY 301 VYVVAERGGDRRLPAVAHVSITLREEAGAYAPLTAHGTLINRVLASCYAVIEEHSMAH 360  
 DB 301 VYVVAERGGDRRLPAVAHVSITLREEAGAYAPLTAHGTLINRVLASCYAVIEEHSMAH 360  
 QY 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWYSQLVHIGTW 420  
 DB 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWYSQLVHIGTW 420  
 QY 421 LIDSETMHPLGMAVKSS 437  
 DB 421 LIDSETMHPLGMAVKSS 437

RESULT 11  
 US-09-057-860A-6  
 Sequence 6, Application US/09057860A  
 Patent No. 6277820  
 GENERAL INFORMATION:  
 APPLICANT: Arnon Rosenthal  
 APPLICANT: Mary Hynes  
 APPLICANT: Weilan Ye  
 TITLE OF INVENTION: Method Of Dopaminergic And Serotonergic  
 TITLE OF INVENTION: Neuron Formation From Neuroprogenitor Cells  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 1 DNA Way  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Winpatin (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/057, 860A  
 FILING DATE: 09-Apr-1998  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Svoboda, Craig G.  
 REGISTRATION NUMBER: 39,044  
 REFERENCE/DOCKET NUMBER: P1364  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650/225-1489  
 TELEFAX: 650/952-9881  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 437 amino acids  
 TYPE: Amino Acid  
 TOPOLOGY: linear  
 US-09-057-860A-6

Query Match 99.9%; Score 2286; DB 4; Length 437;  
 Best Local Similarity 99.8%; Pred. No. 5.9e-249;  
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLARCFVLIVLASSLVCPGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTIGAS 60  
 DB 1 MLLLARCFVLIVLASSLVCPGLACGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTIGAS 60  
 QY 61 GRYEKGITRNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNLAISVNMWPG 120  
 DB 61 GRYEKGITRNSERFKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNLAISVNMWPG 120  
 QY 121 VRLRTGEGDEGHHSESLHTEGRAVDITTSDRDSKYGMLARLAVEGFDMVYTESKA 180  
 DB 121 VRLRTGEGDEGHHSESLHTEGRAVDITTSDRDSKYGMLARLAVEGFDMVYTESKA 180  
 QY 181 HIHCVKAENSVAAKSGCGFPGSATVHLEQGTKLVDLPDGRVLAADDOGRLLYSDFL 240  
 DB 181 HIHCVKAENSVAAKSGCGFPGSATVHLEQGTKLVDLPDGRVLAADDOGRLLYSDFL 240  
 QY 241 TELDRDEGAKKVFYETLEPERELLTPAHLLEFVAPHNDSGPTGPSALFASRVAPGQR 300  
 DB 241 TELDRDEGAKKVFYETLEPERELLTPAHLLEFVAPHNDSGPTGPSALFASRVAPGQR 300  
 QY 301 VYVVAERGGDRRLPAVAHVSITLREEAGAYAPLTAHGTLINRVLASCYAVIEEHSMAH 360  
 DB 301 VYVVAERGGDRRLPAVAHVSITLREEAGAYAPLTAHGTLINRVLASCYAVIEEHSMAH 360  
 QY 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWYSQLVHIGTW 420  
 DB 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPAGIHWYSQLVHIGTW 420  
 QY 421 LIDSETMHPLGMAVKSS 437  
 DB 421 LIDSETMHPLGMAVKSS 437

RESULT 12  
 US-08-349-498-20  
 Sequence 20, Application US/08349498  
 Patent No. 6281332  
 GENERAL INFORMATION:  
 APPLICANT: Beachy, Phillip A.  
 APPLICANT: Moon, Randall T.  
 TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES  
 NUMBER OF SEQUENCES: 20  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fish & Richardson P.C.  
 STREET: 4225 Executive Square, Suite 1400  
 CITY: LA Jolla  
 STATE: CA  
 COUNTRY: U.S.A.

```

; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,498
; FILING DATE: 02-DEC-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/043001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-349-498-20

```

```

Query Match          99.9%; Score 2286; DB 4; Length 437;
Best Local Similarity 99.8%; Pred. No. 5.9e-249;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLVCPGLACGPGRGFGRRPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLIARCFVLIASSLVCPGLACGPGRGFGRRPKLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GREGKTRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATSVNMQWG 120
DB 61 GREGKTRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATSVNMQWG 120
QY 121 VRLRVTEGWDEDEGHSESLHTEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYESKA 180
DB 121 VRLRVTEGWDEDEGHSESLHTEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYESKA 180
QY 181 HICSVKAENSVAAKSGGCPGSAIVHLEOGGTVKLDLRGDRVLAADDOGRLLYSDFL 240
DB 181 HICSVKAENSVAAKSGGCPGSAIVHLEOGGTVKLDLRGDRVLAADDOGRLLYSDFL 240
QY 241 TFLDRDEGAKKVFYIETLEPERILLTAAHLTFVAPHNDSGPTGPGSALFASVRPGR 300
DB 241 TFLDRDEGAKKVFYIETLEPERILLTAAHLTFVAPHNDSGPTGPGSALFASVRPGR 300
QY 301 VYVAERGGDRRLPAAVHSTVTLREEAGAYAPLTAHGTILINRVLASCYAVIEESHMAH 360
DB 301 VYVAERGGDRRLPAAVHSTVTLREEAGAYAPLTAHGTILINRVLASCYAVIEESHMAH 360
QY 361 RAFAFRLAHLAALAPARTDGGGSIIPAASATBARGAETAGIHWYSQLYHIGTW 420
DB 361 RAFAFRLAHLAALAPARTDGGGSIIPAASATBARGAETAGIHWYSQLYHIGTW 420
QY 421 LLDSETHMPLGMAYKSS 437
DB 421 LLDSETHMPLGMAYKSS 437

```

```

RESULT 13
US-09-293-505-14
; Sequence 14, Application US/09293505
; Patent No. 6348575
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/293,505
; CURRENT FILING DATE: 1999-04-15

```

```

; EARLIER APPLICATION NUMBER: US 60/081,884
; EARLIER FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 14
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-293-505-14

```

```

Query Match          99.9%; Score 2286; DB 4; Length 437;
Best Local Similarity 99.8%; Pred. No. 5.9e-249;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLVCPGLACGPGRGFGRRPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLIARCFVLIASSLVCPGLACGPGRGFGRRPKLTPLAYKQFIPNVAEKTIGAS 60
QY 61 GREGKTRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATSVNMQWG 120
DB 61 GREGKTRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATSVNMQWG 120
QY 121 VRLRVTEGWDEDEGHSESLHTEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYESKA 180
DB 121 VRLRVTEGWDEDEGHSESLHTEGRAVDITTSRDRSKYGMRLARLAVEGFDMVYESKA 180
QY 181 HICSVKAENSVAAKSGGCPGSAIVHLEOGGTVKLDLRGDRVLAADDOGRLLYSDFL 240
DB 181 HICSVKAENSVAAKSGGCPGSAIVHLEOGGTVKLDLRGDRVLAADDOGRLLYSDFL 240
QY 241 TFLDRDEGAKKVFYIETLEPERILLTAAHLTFVAPHNDSGPTGPGSALFASVRPGR 300
DB 241 TFLDRDEGAKKVFYIETLEPERILLTAAHLTFVAPHNDSGPTGPGSALFASVRPGR 300
QY 301 VYVAERGGDRRLPAAVHSTVTLREEAGAYAPLTAHGTILINRVLASCYAVIEESHMAH 360
DB 301 VYVAERGGDRRLPAAVHSTVTLREEAGAYAPLTAHGTILINRVLASCYAVIEESHMAH 360
QY 361 RAFAFRLAHLAALAPARTDGGGSIIPAASATBARGAETAGIHWYSQLYHIGTW 420
DB 361 RAFAFRLAHLAALAPARTDGGGSIIPAASATBARGAETAGIHWYSQLYHIGTW 420
QY 421 LLDSETHMPLGMAYKSS 437
DB 421 LLDSETHMPLGMAYKSS 437

```

```

RESULT 14
PCT-US95-15463-20
; Sequence 20, Application PC/TUS9515463
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15463
; FILING DATE: 01-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/080W01

```



TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-15463-20

Query Match 99.9%; Score 2286; DB 5; Length 437;  
Best Local Similarity 99.8%; Pred. No. 5.9e-249;  
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIYASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKOFIPNVAEKTIGAS 60  
DB 1 MLLIARCFVLIYASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKOFIPNVAEKTIGAS 60  
QY 61 GRYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAISVNNQMPG 120  
DB 61 GRYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAISVNNQMPG 120  
QY 121 VRLAVTEGMDGDGHSESLHYEGRAVDITTSDDRSKYGMARLAVAGFDWYTESKA 180  
DB 121 VRLAVTEGMDGDGHSESLHYEGRAVDITTSDDRSKYGMARLAVAGFDWYTESKA 180  
QY 181 HIHCSVAENSVAAKSGCGPFSATVHLEGGTKLVKDLRPGDRVLAADDGRLLYSDFL 240  
DB 181 HIHCSVAENSVAAKSGCGPFSATVHLEGGTKLVKDLRPGDRVLAADDGRLLYSDFL 240  
QY 241 TFLDRDEGAKKVFYIETLEPERELLTPAHLLEFVAPHNDSGPTPGSPALFASRYRPGQR 300  
DB 241 TFLDRDEGAKKVFYIETLEPERELLTPAHLLEFVAPHNDSGPTPGSPALFASRYRPGQR 300  
QY 301 VYVVAERGGDRRLPAAVHSTLREEGAGAYAPLTAHGTIILINRYLASCYAVIEESHMAH 360  
DB 301 VYVVAERGGDRRLPAAVHSTLREEGAGAYAPLTAHGTIILINRYLASCYAVIEESHMAH 360  
QY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIMYSOLLYHIGTW 420  
DB 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIMYSOLLYHIGTW 420  
QY 421 LLDSETHPLGMAVKSS 437  
DB 421 LLDSETHPLGMAVKSS 437

RESULT 15  
PCT-US95-15923-20  
Sequence 20, Application PC/TUS9515923  
GENERAL INFORMATION:  
APPLICANT: The Johns Hopkins University School of Medicine, et al.  
TITLE OF INVENTION: NOVEL HEDGEGHOG-DERIVED POLYPEPTIDES  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/15923  
FILING DATE: 04-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.

REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/043W01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-15923-20

Query Match 99.9%; Score 2286; DB 5; Length 437;  
Best Local Similarity 99.8%; Pred. No. 5.9e-249;  
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIYASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKOFIPNVAEKTIGAS 60  
DB 1 MLLIARCFVLIYASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKOFIPNVAEKTIGAS 60  
QY 61 GRYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAISVNNQMPG 120  
DB 61 GRYEGKITRNSERKELTPNPNPDIIFKDEENTGADRLMTQRCCKLNALAISVNNQMPG 120  
QY 121 VRLAVTEGMDGDGHSESLHYEGRAVDITTSDDRSKYGMARLAVAGFDWYTESKA 180  
DB 121 VRLAVTEGMDGDGHSESLHYEGRAVDITTSDDRSKYGMARLAVAGFDWYTESKA 180  
QY 181 HIHCSVAENSVAAKSGCGPFSATVHLEGGTKLVKDLRPGDRVLAADDGRLLYSDFL 240  
DB 181 HIHCSVAENSVAAKSGCGPFSATVHLEGGTKLVKDLRPGDRVLAADDGRLLYSDFL 240  
QY 241 TFLDRDEGAKKVFYIETLEPERELLTPAHLLEFVAPHNDSGPTPGSPALFASRYRPGQR 300  
DB 241 TFLDRDEGAKKVFYIETLEPERELLTPAHLLEFVAPHNDSGPTPGSPALFASRYRPGQR 300  
QY 301 VYVVAERGGDRRLPAAVHSTLREEGAGAYAPLTAHGTIILINRYLASCYAVIEESHMAH 360  
DB 301 VYVVAERGGDRRLPAAVHSTLREEGAGAYAPLTAHGTIILINRYLASCYAVIEESHMAH 360  
QY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIMYSOLLYHIGTW 420  
DB 361 RAFAFRLAHLAALAPARTDGGGGSIPAAQSATEARGAEPAGIMYSOLLYHIGTW 420  
QY 421 LLDSETHPLGMAVKSS 437  
DB 421 LLDSETHPLGMAVKSS 437

Search completed: February 20, 2003, 10:14:11  
Job time : 12.4398 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:08:18 ; Search time 14.3815 seconds  
(without alignments)  
2921.176 Million cell updates/sec

Title: US-09-827-110a-13

Perfect score: 2289

Sequence: 1 MLLILARCFVLVIASSLLVC.....GTWLDSETHMPLGNVAKSS 437

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2286	99.9	437	2	A49425
2	2200	96.1	437	2	B53193
3	1788	78.1	425	2	A49424
4	1525.5	66.6	444	2	S56765
5	1517	66.3	418	2	A53193
6	1458	63.7	415	2	A49426
7	1219	53.3	396	2	B49425
8	1043	45.6	336	2	C49425
9	894	39.1	471	2	A46400
10	452	19.7	94	2	G02735
11	170	7.4	615	2	T29550
12	153.5	6.7	1207	2	T23754
13	148	6.5	1226	2	T24045
14	146	6.4	481	2	T27655
15	140.5	6.1	629	2	T19563
16	130.5	5.7	868	2	T22281
17	128	5.6	1021	2	T23252
18	127	5.5	484	2	T34504
19	118	5.2	598	2	T42070
20	117	5.1	1407	2	T28702
21	108	4.7	481	2	T27975
22	107	4.7	672	2	T36083
23	105	4.6	1324	2	T14070
24	102.5	4.5	384	2	H87019
25	101	4.4	482	2	E70554
26	100.5	4.4	863	2	H84177
27	100.5	4.4	930	2	T35180
28	100	4.4	831	2	H84368
29	100	4.4	1147	2	D87295

30	99	4.3	869	2	S76720	hypothetical prote
31	98.5	4.3	532	2	E87343	ABC transporter, A
32	98.5	4.3	1464	2	T13716	bazooka gene prote
33	98	4.3	940	2	E87250	[protein-pil] urid
34	97.5	4.3	313	2	T15855	hypothetical prote
35	96	4.2	205	2	T26220	hypothetical prote
36	96	4.2	373	2	D36067	thyroid hormone re
37	96	4.2	419	2	AF1274	retrovirus-related
38	96	4.2	572	1	FOH415	copper resistance
39	96	4.2	611	2	C82845	probable exoribonu
40	96	4.2	759	2	G71279	hypothetical prote
41	96	4.2	1958	2	B40505	probable outer mem
42	95.5	4.2	1531	2	H71468	gene c-erbA-beta p
43	95	4.2	373	2	T51165	probable anthranil
44	94.5	4.1	335	2	T36304	proplyl-CoA card
45	94.5	4.1	669	2	B87519	

#### ALIGNMENTS

##### RESULT 1

A49425 Sonic hedgehog protein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999

C:Accession: A49425

R:Rehder, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M

Cell 75, 1417-1430, 1993

A:Title: Sonic hedgehog, a member of a family of putative signalling molecules, is imp

A:Reference number: A49425; MUID:94094334; PMID:7916661

A:Accession: A49425

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1437 <ECH>

A:Cross-references: GB:X76290

A>Note: authors translation is shown for the codon TCC at position 436

C:Genetics:

A:Gene: Shh

C:Superfamily: sonic hedgehog protein

Query Match	99.9%	Score 2286	DB 2	Length 437
Best Local Similarity	99.8%	Pred. No. 7.2e-177		
Matches 436	Conservative 1	Mismatches 0	Indels 0	Gaps 0
QY 1	MLLILARCFVLVIASSLLVC	PGIACG	RGGRGRKRHKITPLAYKQPIFVNAKTI	IGAS 60
DB 1	MLLILARCFVLVIASSLLVC	PGIACG	RGGRGRKRHKITPLAYKQPIFVNAKTI	IGAS 60
QY 61	GRYGGKTRNSERPKELTPNYNPDIIFKDEENTGADRLMTORCKDKLNALAI	SVNOMP	60	
DB 61	GRYGGKTRNSERPKELTPNYNPDIIFKDEENTGADRLMTORCKDKLNALAI	SVNOMP	60	
QY 121	VRLAVTEGWDGDGHSESLHYEGRAVDITSDRDRSKYGLARLAVEAGFDWVYESKA	180		
DB 121	VRLAVTEGWDGDGHSESLHYEGRAVDITSDRDRSKYGLARLAVEAGFDWVYESKA	180		
QY 181	HIRHVSVAENSVAKSGCGPGSATVHL	EOGCTLVLDLRGDMVLAADQGRLLYSDFL	240	
DB 181	HIRHVSVAENSVAKSGCGPGSATVHL	EOGCTLVLDLRGDMVLAADQGRLLYSDFL	240	
QY 241	TFILDRDCAKRVFVITLPERRELLTAAHLFVA	RHNSGPTPGSALFASVRGGR	300	
DB 241	TFILDRDCAKRVFVITLPERRELLTAAHLFVA	RHNSGPTPGSALFASVRGGR	300	
QY 301	VYVVAERGGDRLLPAAVHSVTLREEDAGAVPTL	AGHTILINVLASCAVAIEHSMH	360	
DB 301	VYVVAERGGDRLLPAAVHSVTLREEDAGAVPTL	AGHTILINVLASCAVAIEHSMH	360	
QY 361	RAFAPEFLAHALLAALPARTDGGGGG	SIPAAGSATRAEAPAGIHW	SOLLVYHGTW	420
DB 361	RAFAPEFLAHALLAALPARTDGGGGG	SIPAAGSATRAEAPAGIHW	SOLLVYHGTW	420

QY 421 LLDSETHPLGMAVKSS 437  
 |||||  
 Db 421 LLDSETHPLGMAVKSS 437

## RESULT 2

Best Local Similarity 96.1%; Score 2200; DB 2; Length 437;  
 Matches 421; Conservative 3; Mismatches 13; Indels 0; Gaps 0;  
 hedgehog homolog vhh-1 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999  
 C:Accession: B53193  
 R:Roelink, H.; Augburger, A.; Heemskerk, J.; Korzh, V.; Norlin, S.; Ruiz i Altaba, A.;  
 Cell 76, 761-775, 1994  
 A:Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgeh  
 A:Reference number: A53193; MID:94170375; PMID:8124714  
 A:Accession: B53193  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-437 <ROE>  
 A:Cross-references: GB:L27340; MID:9452122; PIDN:AAA20999.1; PID:9452123  
 C:Superfamily: sonic hedgehog protein

Query Match 96.1%; Score 2200; DB 2; Length 437;  
 Best Local Similarity 96.3%; Pred. No. 6.5e-170;  
 Matches 421; Conservative 3; Mismatches 13; Indels 0; Gaps 0;  
 QY 1 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 60  
 |||||  
 Db 1 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 60  
 QY 61 GYEGKITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAIISVNQMPG 120  
 |||||  
 Db 61 GYEGKITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAIISVNQMPG 120  
 QY 121 VRLRVTEGDEDEGHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYESKA 180  
 |||||  
 Db 121 VRLRVTEGDEDEGHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYESKA 180  
 QY 121 VRLRVTEGDEDEGHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYESKA 180  
 |||||  
 Db 121 VRLRVTEGDEDEGHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYESKA 180  
 QY 181 RHCSYKAENSVAAKSGCGPSATVHLEOGTKLVKDLRPDRLAADDOGRLLYSDFL 240  
 |||||  
 Db 181 RHCSYKAENSVAAKSGCGPSATVHLEOGTKLVKDLRPDRLAADDOGRLLYSDFL 240  
 QY 241 TELDRREGAKKVFYETLEPRERLITLTAHLTFVAPRHNDSCPTPPSALFASRVPRQR 300  
 |||||  
 Db 241 TELDRREGAKKVFYETLEPRERLITLTAHLTFVAPRHNDSCPTPPSALFASRVPRQR 300  
 QY 301 VVVAERGGDRRLTPAAVHVTLREBEAGAVAPLTAHGFTILINRVLASCYAVIEESHMAH 360  
 |||||  
 Db 301 VVVAERGGDRRLTPAAVHVTLREBEAGAVAPLTAHGFTILINRVLASCYAVIEESHMAH 360  
 QY 361 RNFAPRRLAHLALALAPARTDGGGGSIPAAQSATEARGAPETAGIHWSQLYHIGTW 420  
 |||||  
 Db 361 RNFAPRRLAHLALALAPARTDGGGGSIPAAQSATEARGAPETAGIHWSQLYHIGTW 420  
 QY 421 LLDSETHPLGMAVKSS 437  
 |||||  
 Db 421 LLDSETHPLGMAVKSS 437

## RESULT 3

Best Local Similarity 67.4%; Pred. No. 2.1e-115;  
 Matches 304; Conservative 38; Mismatches 88; Indels 21; Gaps 6;  
 patterning protein Sonic hedgehog precursor - chicken  
 C:Species: Gallus gallus (chicken)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 21-Jul-2000  
 C:Accession: A49424  
 R:Riddle, R.D.; Johnson, R.L.; Lauffer, E.; Tabin, C.  
 Cell 75, 1401-1416, 1993  
 A:Title: Sonic hedgehog mediates the polarizing activity of the ZPA.  
 A:Reference number: A49424; MID:94094333; PMID:8269518  
 A:Accession: A49424  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-425 <RID>

A:Cross-references: GB:L28099; MID:9453526; PIDN:AAA72428.1; PID:9453527  
 C:Superfamily: sonic hedgehog protein  
 F:1-26/Domain: signal sequence #status predicted <SIG>

Query Match 78.1%; Score 1788; DB 2; Length 425;  
 Best Local Similarity 80.2%; Pred. No. 1.2e-136;  
 Matches 353; Conservative 22; Mismatches 43; Indels 22; Gaps 6;

QY 2 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 61  
 |||||  
 Db 4 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 63  
 QY 62 RYEGKITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAIISVNQMPG 121  
 |||||  
 Db 62 RYEGKITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAIISVNQMPG 123  
 QY 122 RLRVTEGDEDEGHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYESKAH 181  
 |||||  
 Db 124 KLRVTEGDEDEGHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYESKAH 183  
 QY 182 RHCSYKAENSVAAKSGCGPSATVHLEOGTKLVKDLRPDRLAADDOGRLLYSDFL 241  
 |||||  
 Db 184 RHCSYKAENSVAAKSGCGPSATVHLEOGTKLVKDLRPDRLAADDOGRLLYSDFL 243  
 QY 242 FLDRREGAKKVFYETLEPRERLITLTAHLTFVAPRHNDSCPTPPSALFASRVPR 297  
 |||||  
 Db 244 FLDRREGAKKVFYETLEPRERLITLTAHLTFVAPRHNDSCPTPPSALFASRVPR 302  
 QY 298 GGRVYVAERGGDRRLTPAAVHVTLREBEAGAVAPLTAHGFTILINRVLASCYAVIEES 357  
 |||||  
 Db 303 GGRVYVLEGE--GGQQLPRAVSHSVSLREBSAGVAPLTAHGFTILINRVLASCYAVIEES 360  
 QY 358 WNRAPRRLAHLALALAPARTDGGGGSIPAAQSATEARGAPETAGIHWSQLYHIGTW 417  
 |||||  
 Db 361 WNRAPRRLAHLALALAPARTDGGGGSIPAAQSATEARGAPETAGIHWSQLYHIGTW 405  
 QY 418 GWLDDSETHPLGMAVKSS 437  
 |||||  
 Db 406 GSWVLDGDLHPLGMAVPAS 425

## RESULT 4

Best Local Similarity 66.6%; Score 1525.5; DB 2; Length 444;  
 Matches 304; Conservative 38; Mismatches 88; Indels 21; Gaps 6;  
 morphogen Xhh precursor - African clawed frog  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 10-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 21-Jul-2000  
 C:Accession: S56765  
 R:Stolow, M.A.; Shi, Y.B.  
 Nucleic Acids Res. 23, 2555-2562, 1995  
 A:Title: Xenopus sonic hedgehog as a potential morphogen during embryogenesis and thy  
 A:Reference number: S56765; MID:95357169; PMID:7630736  
 A:Accession: S56765  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-444 <STD>  
 A:Cross-references: EMBL:L39213; MID:9790937; PIDN:AA42227.1; PID:9790938  
 C:Superfamily: sonic hedgehog protein

Query Match 66.6%; Score 1525.5; DB 2; Length 444;  
 Best Local Similarity 67.4%; Pred. No. 2.1e-115;  
 Matches 304; Conservative 38; Mismatches 88; Indels 21; Gaps 6;  
 QY 1 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 60  
 |||||  
 Db 1 MLLLAFCFLVLIASSLVCPGLACGPGRGFRHHPKRLTPLAYKQFIPNVAEKTIGAS 60  
 QY 61 GYEGKITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAIISVNQMPG 120  
 |||||  
 Db 61 GYEGKITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDKLNALAIISVNQMPG 120  
 QY 121 VRLRVTEGDEDEGHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYESKA 180  
 |||||  
 Db 121 VRLRVTEGDEDEGHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDMVYESKA 180

OY 181 HHCISVKAENSYAAGSGCGFPGSATVHLEGGGKLVKDLRPGDRVLADDOGRLLTDFL 240  
 DB 181 HHCISVKAENSYAAGSGCGFPGSATVHLEGGGKLVKDLRPGDRVLADDOGRLLTDFL 240  
 OY 241 TFLDRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300  
 DB 241 TFLDRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300  
 OY 301 VVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILLINRVLASCYAIEEHSMAHRA 360  
 DB 301 VVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILLINRVLASCYAIEEHSMAHRA 360  
 OY 361 RAFAERLHAALLAALAPARTDGGGGSIPAAO-----SATEARGAEPRTAG 406  
 DB 361 RAFAERLHAALLAALAPARTDGGGGSIPAAO-----SATEARGAEPRTAG 406  
 OY 407 IHWSQLYHIGTWLSDSETMHPGLMAVKS 437  
 DB 407 IHWSQLYHIGTWLSDSETMHPGLMAVKS 437  
 OY 414 IHWSQLYHIGTWLSDSETMHPGLMAVKS 444  
 DB 414 IHWSQLYHIGTWLSDSETMHPGLMAVKS 444

## RESULT 5

A:Species: Brachydanio rerio (zebra fish)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999  
 C:Accession: A53193  
 R:Roelink, H.; Augsburger, A.; Heemskerk, J.; Korzh, V.; Norlin, S.; Ruiz, I. Altaba, A.;  
 Cell 76, 761-775, 1994  
 A:Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgehog  
 A:Reference number: A53193; MUID:94170375; PMID:8124714  
 A:Accession: A53193  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-418 <KRA>  
 A:Cross-references: GB:L27585; NID:9452159; PID:9452160  
 C:Superfamily: sonic hedgehog protein

Query Match 66.3%; Score 1517; DB 2; Length 418;  
Best Local Similarity 67.7%; Pred. No. 9, 6e-115;

Matches 294; Conservative 47; Mismatches 75; Indels 18; Gaps 3;

OY 4 LARCLVLIASSLVCPGLACGPGRGFRKRRHKKLTPLAYKQFIPTNAEKTLCASGRY 63  
 DB 4 LARCLVLIASSLVCPGLACGPGRGFRKRRHKKLTPLAYKQFIPTNAEKTLCASGRY 63  
 OY 3 LTRVLVLSLTLVSLVSGIACGPGRGYRRHKKLTPLAYKQFIPTNAEKTLCASGRY 62  
 DB 3 LTRVLVLSLTLVSLVSGIACGPGRGYRRHKKLTPLAYKQFIPTNAEKTLCASGRY 62  
 OY 64 EGRITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDLNLAISVNMQPGVRL 123  
 DB 64 EGRITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDLNLAISVNMQPGVRL 123  
 OY 63 EGRITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDLNLAISVNMQPGVRL 122  
 DB 63 EGRITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDLNLAISVNMQPGVRL 122  
 OY 124 RTEGDEDEGHHSESLHTEGRAVDITTSDRDSKYGMRLAVALVEGPMVYVESKAH 183  
 DB 124 RTEGDEDEGHHSESLHTEGRAVDITTSDRDSKYGMRLAVALVEGPMVYVESKAH 183  
 OY 123 RTEGDEDEGHHSESLHTEGRAVDITTSDRDSKYGMRLAVALVEGPMVYVESKAH 182  
 DB 123 RTEGDEDEGHHSESLHTEGRAVDITTSDRDSKYGMRLAVALVEGPMVYVESKAH 182  
 OY 184 CSYKAENSYAAGSGCGFPGSATVHLEGGGKLVKDLRPGDRVLADDOGRLLTDFL 243  
 DB 184 CSYKAENSYAAGSGCGFPGSATVHLEGGGKLVKDLRPGDRVLADDOGRLLTDFL 243  
 OY 183 CSYKAENSYAAGSGCGFPGSATVHLEGGGKLVKDLRPGDRVLADDOGRLLTDFL 242  
 DB 183 CSYKAENSYAAGSGCGFPGSATVHLEGGGKLVKDLRPGDRVLADDOGRLLTDFL 242  
 OY 244 DRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 303  
 DB 244 DRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 303  
 OY 304 VVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILLINRVLASCYAIEEHSMAHRA 363  
 DB 304 VVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILLINRVLASCYAIEEHSMAHRA 363  
 OY 364 RAFAERLHAALLAALAPARTDGGGGSIPAAO-----SATEARGAEPRTAG 406  
 DB 364 RAFAERLHAALLAALAPARTDGGGGSIPAAO-----SATEARGAEPRTAG 406  
 OY 407 IHWSQLYHIGTWLSDSETMHPGLMAVKS 437  
 DB 407 IHWSQLYHIGTWLSDSETMHPGLMAVKS 437

DB 405 SNMLHPLGMSVNS 418

## RESULT 6

A:Species: Brachydanio rerio (zebra fish)  
 C:Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 10-Dec-1999  
 C:Accession: A49426  
 R:Krauss, S.; Concordet, J.P.; Ingham, P.W.  
 Cell 75, 1431-1444, 1993  
 A:Title: A functionally conserved homolog of the Drosophila segment polarity gene hh  
 A:Reference number: A49426; MUID:94094335; PMID:826519  
 A:Accession: A49426  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-415 <KRA>  
 A:Experimental source: embryo  
 A:Note: sequence extracted from NCBI backbone (NCBIP:142459)  
 C:Superfamily: sonic hedgehog protein

Query Match 63.7%; Score 1458; DB 2; Length 415;  
Best Local Similarity 67.1%; Pred. No. 5, 6e-110;

Matches 289; Conservative 46; Mismatches 78; Indels 18; Gaps 5;

OY 4 LARCLVLIASSLVCPGLACGPGRGFRKRRHKKLTPLAYKQFIPTNAEKTLCASGRY 63  
 DB 4 LARCLVLIASSLVCPGLACGPGRGFRKRRHKKLTPLAYKQFIPTNAEKTLCASGRY 63  
 OY 3 LTRVLVLSLTLVSLVSGIACGPGRGYRRHKKLTPLAYKQFIPTNAEKTLCASGRY 62  
 DB 3 LTRVLVLSLTLVSLVSGIACGPGRGYRRHKKLTPLAYKQFIPTNAEKTLCASGRY 62  
 OY 64 EGRITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDLNLAISVNMQPGVRL 123  
 DB 64 EGRITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDLNLAISVNMQPGVRL 123  
 OY 63 EGRITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDLNLAISVNMQPGVRL 122  
 DB 63 EGRITNSERFELTPNYPDIIFKDEENTGADRLMTORCKDLNLAISVNMQPGVRL 122  
 OY 124 RTEGDEDEGHHSESLHTEGRAVDITTSDRDSKYGMRLAVALVEGPMVYVESKAH 183  
 DB 124 RTEGDEDEGHHSESLHTEGRAVDITTSDRDSKYGMRLAVALVEGPMVYVESKAH 183  
 OY 123 RTEGDEDEGHHSESLHTEGRAVDITTSDRDSKYGMRLAVALVEGPMVYVESKAH 182  
 DB 123 RTEGDEDEGHHSESLHTEGRAVDITTSDRDSKYGMRLAVALVEGPMVYVESKAH 182  
 OY 184 CSYKAENSYAAGSGCGFPGSATVHLEGGGKLVKDLRPGDRVLADDOGRLLTDFL 243  
 DB 184 CSYKAENSYAAGSGCGFPGSATVHLEGGGKLVKDLRPGDRVLADDOGRLLTDFL 243  
 OY 183 CSYKAENSYAAGSGCGFPGSATVHLEGGGKLVKDLRPGDRVLADDOGRLLTDFL 242  
 DB 183 CSYKAENSYAAGSGCGFPGSATVHLEGGGKLVKDLRPGDRVLADDOGRLLTDFL 242  
 OY 244 DRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 303  
 DB 244 DRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 303  
 OY 243 DRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 301  
 DB 243 DRDEGAKKVFYIETLEPERELLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 301  
 OY 304 VVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILLINRVLASCYAIEEHSMAHRA 363  
 DB 304 VVVAERGGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILLINRVLASCYAIEEHSMAHRA 363  
 OY 364 RAFAERLHAALLAALAPARTDGGGGSIPAAO-----SATEARGAEPRTAG 406  
 DB 364 RAFAERLHAALLAALAPARTDGGGGSIPAAO-----SATEARGAEPRTAG 406  
 OY 407 IHWSQLYHIGTWLSDSETMHPGLMAVKS 434  
 DB 407 IHWSQLYHIGTWLSDSETMHPGLMAVKS 434

## RESULT 7

A:Species: Mus musculus (house mouse)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999  
 C:Accession: B49425  
 R:Rechardt, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M  
 Cell 75, 1417-1430, 1993  
 A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is imp  
 A:Reference number: B49425; MUID:94094344; PMID:7916661  
 A:Accession: B49425  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-396 <ECH>

A:Cross-references: GB:X76292; NID:g443941; PIDN:CAA53924.1; PID:g443942  
 C:Genetics:  
 A:Gene: Dhh  
 C:Superfamily: sonic hedgehog protein

Query Match 53.3%; Score 1219; DB 2; Length 396;  
 Best Local Similarity 58.7%; Pred. No. 1.1e-90;  
 Matches 249; Conservative 55; Mismatches 80; Indels 40; Gaps 12;

OY 4 LLARCFVLIVLSSLLVCGLAGCGRG-FGKRHRPK-LTPAYKQIFPNVAKTLGASG 61  
 DB 7 LRLPLCLALLALS-----AOSGPGRGVRRRYRKQVLPYKQFSPMERLIGASG 61  
 OY 62 REGKITRNSERFKELTPNYNDITFKDEBNTGADRLMTQCKDKLNAISVNNQMPGV 121  
 DB 62 PAEGVTRKSERFRDLVNYNDITFKDEBNSGADRLMTCKERVNALALAVNNMMPGV 121  
 OY 122 RLRTVEGDEDDHSEESLHVEGRAVDITTSDRDRSKYGMRLAVEAGFDWVYESKAH 181  
 DB 122 RLRTVEGDEDDHMODSLHVEGRALDITTSDRDRNKIGLARLAVEAGFDWVYESRKH 181  
 OY 182 IHCYKAENSVAAKSGCGFPGSATVHLBOGCTKLVKDLRPGDRLVLAADQGRILYSDEL 241  
 DB 182 IHVYKADNSLAVRAGCGFPGNATVRLRSGERKGLRELHGDWVLAADAGVPTPLVLL 241  
 OY 242 FLDRBEGAKVYVYETLEPRERLLITAAHLIFVAPHNDSGPTGP-----SALFASRVBP 298  
 DB 242 FLDRLOQRASFAVETERPRPKLLITPMHLYFAA-----RGPAPAPGDFAPVFAHRLRAG 297  
 OY 299 QRVYVVARSGRRLLPAVHSVTLREBAGAVAPLTAHGTLLINRYLASCYAVIEESHW 358  
 DB 298 DSVLA---PGGD-ALQPARVAVR-REBAVGAFPLTAHGTLLVNDVLAESCVAIESHOW 352  
 OY 359 AHRAAPRLAHLALAPARTDGGGSGIPAAQSATEARGAEPTAGIHWSQLYHIG 418  
 DB 353 AHRAAPRLHLHA-LGALLP-----GGAV-----QPT-GMHWSRLYLA 391  
 OY 419 TWLL 422  
 DB 392 EELM 395

## RESULT 8

C49425  
 Indian hedgehog protein - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 11-Jan-2000  
 C:Accession: C49425  
 R:Exelard, Y.; Egestin, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; McMahon, J.A.; Title: Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the development of the mouse embryo. *Cell* 75, 1417-1430, 1993  
 A:Reference number: A49425; MUID:94094334; PMID:7916661  
 A:Accession: C49425  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-336 <ECH>  
 A:Cross-references: GB:X76291  
 C:Superfamily: sonic hedgehog protein

Query Match 45.6%; Score 1043; DB 2; Length 336;  
 Best Local Similarity 58.7%; Pred. No. 1.5e-78;  
 Matches 213; Conservative 36; Mismatches 80; Indels 32; Gaps 5;  
 OY 72 ERFKELPNYNDITFKDEBNTGADRLMTQCKDKLNAISVNNQMPGVRLRTVEGDE 131  
 DB 1 ERFKELPNYNDITFKDEBNTGADRLMTQCKDKRLNSLAISVNNQMPGVRLRTVEGDE 60  
 OY 132 DGHSEESLHVEGRAVDITTSDRDRSKYGMRLAVEAGFDWVYESKAHITCSVKAENS 191  
 DB 61 DGHSEESLHVEGRAVDITTSDRDRNKYGLARLAVEAGFDWVYESKAHVCSVKSBS 120  
 OY 192 VAAKSGCFPGSATVHLBOGCTKLVKDLRPGDRLVLAADQGRILYSDELTLDRDEGAKK 251  
 DB 192 VAAKSGCFPGSATVHLBOGCTKLVKDLRPGDRLVLAADQGRILYSDELTLDRDEGAKK 251

DB 121 AAKTGCGFPAAGVRLNENGERVALSAVKPGDRLVLAAGEDTPTFSVLIPLDRPNRLR 180  
 OY 252 VEVYIETLEPRERLLITAAHLIFVAPHNDSGPTGPSALFASRVPGQRYVVAERGGDR 311  
 DB 181 AFQVETQDPPRRRLALTFAHLLFLA-DNHTERPAHFRATFASHQPGQYLVSGYCPG--- 236  
 OY 312 RLPAVHSVTLREBAGAVAPLTAHGTLLINRYLASCYAVIEESHWARAFAPRLAHA 371  
 DB 237 -LQPARVAAS-THVALSYAPLTRHGTLLVEDVVASCFAAVADHNLQLAFWPLRLRPS 294  
 OY 372 LLAALAPARTDGGGSGIPAAQSATEARGA-EPTAGIHWSQLYHIGTLLDSTHPL 430  
 DB 295 L-----AWGSWTPSGVHSYQMLYRLGLRLLESTFHL 329  
 OY 431 GMA 433  
 DB 330 GMS 332

## RESULT 9

A46400  
 segment polarity protein hedgehog - fruit fly (Drosophila melanogaster)

N:Alternate names: hh protein  
 C:Species: Drosophila melanogaster  
 C:Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 10-Dec-1999  
 C:Accession: A46400; JN0501; A43480  
 R:Tabata, T.; Eaton, S.; Kornberg, T.B.  
 Gene Dev. 6, 2635-2645, 1992

A:Title: The Drosophila hedgehog gene is expressed specifically in posterior compartment  
 A:Reference number: A46400; MUID:94040725; PMID:1340474  
 A:Accession: A46400  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-471 <TRAB>  
 A:Cross-references: GB:S66384; NID:g435848; PID:g435849  
 A:Experimental source: Oregon R  
 A:Note: sequence extracted from NCBI backbone (NCBIN:138996, NCBIPI:138997)  
 R:Tashiro, S.; Michie, T.; Higashijima, S.; Zennou, S.; Ishimaru, S.; Takahashi, F.; Gene 124, 183-189, 1993  
 A:Title: Structure and expression of hedgehog, a Drosophila segment-polarity gene req  
 A:Reference number: JN0501; MUID:93185922; PMID:8166882  
 A:Accession: JN0501  
 A:Molecule type: mRNA  
 A:Residues: 1-471 <TRAS>  
 A:Cross-references: GB:I05404  
 A:Note: it is uncertain whether Met-1 or Met-51 is the initiator  
 A:Note: intron positions were determined from partial DNA sequence  
 R:Lee, J.T.; von Kessler, D.P.; Parks, S.; Beachy, P.A.  
 Cell 71, 33-50, 1992  
 A:Title: Secretion and localized transcription suggest a role in positional signaling  
 A:Reference number: A43480; MUID:93008241; PMID:1394430  
 A:Accession: A43480  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-471 <LEB>  
 A:Cross-references: GB:L02793; NID:g157609; PID:g157610  
 A:Note: sequence extracted from NCBI backbone (NCBIPI:115418)  
 C:Comment: This protein is required for cell-cell communication.  
 C:Genetics:  
 A:Gene: hh  
 A:Cross-references: Flybase:FBgn0004644  
 A:Introns: 160/3; 248/1  
 C:Superfamily: sonic hedgehog protein  
 C:Keywords: transmembrane protein  
 F:62-82/Domain: transmembrane #status predicted <TRM>

Query Match 39.1%; Score 894; DB 2; Length 471;  
 Best Local Similarity 46.8%; Pred. No. 2.5e-64;  
 Matches 197; Conservative 62; Mismatches 140; Indels 22; Gaps 9;  
 OY 10 LVLIASSLVPGLAGCGRGFGKRRHPKLTPLAYKQIFPNVAKTLGASGRREGKITR 69  
 DB 70 LLTLVLPVSPHSCGPGRLGRHR-ARNLVPLVLTQIFPNISEYNSASGPLEGVIRR 128

```

Oy 70 NSERKELTPNNNDIIFPEDEENTGADRLMTQRCDKLNALISVNNQMPGYRLRTYEGW 129
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 DSPKFKDLVPNNNDIIFPEDEGTGADRLMSKCKEKLVLVAVNNEMPGRLLVLTESW 188
Oy 130 DEDGHHSSESLYEEERAVDITTSDDRSKYGMLARLAVAGFDWVYVESKAIHCSVRAE 189
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 DEDYHHGGEESLYEEGRVATYATSDSDQSKYGLARLAVAGFDWVSYSRRHITCVSKSD 248
Oy 190 NSVAKSGCGCFPGSATVHLEGGCTLYKVDLRPGDRVLAADDGGRLLYSDFLFLDDEGA 249
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 249 SSISHHVGCFFPESTALLESQVRKPLGELSIGDRVLSMTANGAVYSEVILFMDRLNQ 308
Oy 250 KKVVFVIELERREULLLAHLTRY-APHNDSGPRPGSALFASRYRGQVYVVAEAG 308
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 309 MONFQVLTR-DEGAVLTVYPAHLVSVMQDESK-----LTFPFAIRIEKNQVLVANDVET 362
Oy 309 GDRRLVPAVHSV-TLREEAGAVPLTFHGTILLINRVASCYAVVIEESHMAHRAFAPR 367
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 363 GBLR--PQRVYVGVSGRS--GVVAPLTFEGITVYNSVAASCTAVYINQSLAHMGLAPAR 418
Oy 368 LAHALLAALAPARTDGGGGGSIIPAAQSATANGAEFPACIHWYSOLLVYHIGTWLDSFTM 427
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 419 LLSTLEAWL-PAKEO-----LHSSPRVSSAQOQNGIHMYANALYKVKDYLVEQSWR 469
Oy 428 H 428
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 470 H 470

RESULT 10
G02735
desert hedgehog - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 11-Jan-2000
C:Accession: G02735
R:Drummond, I. A.
submitted to the EMBL Data Library, June 1996
A:Reference number: H01643
A:Accession: G02735
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-94 <DRU>
A:Cross-references: EMBL:U59748; NID:g1401271; PID:g1401272
C:Genetics:
A:Gene: hdnH
C:Superfamily: sonic hedgehog protein

Query Match 19.7%; Score 452; DB 2; Length 94;
Best Local Similarity 85.1%; Pred. No.1.5e-29;
Matches 80; Conservative 12; Mismatches 2; Indels 0; Gaps 0

Oy 85 IIFKDEBNTGADRLMTQRCDKLNALISVNNQMPGVRLVTEGWDGCHSSESLHYEG 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 IIFKDEBNSGADRLMTQRCCKERYNALIAYMMMPGVRLVTEGWDGCHHADSLEYEG 60
Oy 145 RAVDITTSDDRSKYGMLARLAVAGFDWVYVES 178
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 RALDITTSDDRKNTGLARLAVAGFDWVYIGS 94

RESULT 11
T29550
hypocholelcal protein ZK377.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T29550
R:Nhan, M.; Hawkins, J.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid ZK377.
A:Reference number: Z20639
A:Accession: T29550
A:Status: preliminary; translated from GB/EMBL/DBJ

```

[illegible]

Db 1116 SHLESBS-AHKLIIDF 1130

# RESULT 13

724045

hypothetical protein R08B4.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T24045

R:White, S.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19834

A:Accession: T24045

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1226 <Wtl>

A:Cross-references: EMBL:Z68008; PIDN:CA92000.1; GSPDB:GN00028; CESP:R08B4.1

A:Experimental source: clone R08B4

C:Genetics:

A:Gene: CESP:R08B4.1

A:Map position: X

A:Introns: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 839/3; 877/1; 960/3; 1

Query Match

Best Local Similarity 26.8%; Score 148; DB 2; Length 1226;

Matches 51; Conservative 39; Mismatches 66; Indels 34; Gaps 9;

Db 190 NSVAKSGCGPGSATVHLEGGTKLVKDLRPGDRVLAADDOGRLLYSDFLTFLDRDGA 249

Db 1012 SALVATACGCSLDTWV-TTPGKKRMOIDIGYVLADLE-KYFTPTILMIREPEK 1069

QY 250 KKFVYVETLEPRERLLTAHLF-----VAPHNSGPTPGPSALFASRYR 296

Db 1070 VOEFLLTMT-EXGKTLRTSRHFMYRNKCGKSYQYIKMLPHDG-----EAFASDLE 1121

QY 297 PGQRYVAERGGDRLLPAVHSTLREERGAAPLTANGTILINRYLASCYVIEEH 356

Db 1122 VDDCV-VVLYGKRYQ---QKIEITV-RSVRTGYSPLTNCRITVNDMLASCYSEIQON 1176

QY 357 S-----NAH 360

Db 1177 TLQTFEWMAY 1186

# RESULT 14

727665

hypothetical protein ZK1037.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T27665

R:Baslam, V.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z20401

A:Accession: T27665

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-481 <Wtl>

A:Cross-references: EMBL:Z81142; PIDN:CA803509.1; GSPDB:GN00023; CESP:ZK1037.10

A:Experimental source: clone ZK1037

C:Genetics:

A:Gene: CESP:ZK1037.10

A:Map position: 5

A:Introns: 37/3; 75/3; 141/3; 173/2; 253/1; 298/3; 310/1; 371/2; 424/3

Query Match

Best Local Similarity 25.6%; Score 146; DB 2; Length 481;

Matches 43; Conservative 35; Mismatches 70; Indels 20; Gaps 5;

QY 199 CPGSATVHLEGGTKLVKDLRPGDRVLAADDOGRLLYSDFLTFL-----DRDEGAK 250

Db 314 CFPNDAYVNVYEKAVKRDELEIGDWVADENG-----EDITFLPYKWLHROPDEQAE 368

QY 251 KVFYVETLEPRERLLTAHLFVAP--HNSGPTPGPSALFASRYRPGQRYVVAERG 308

Db 369 ---FILESLDNGEFTLTLEKHLVYTECRONSSELKISMESISACKVNAOCFYLAQSEA 425

QY 309 GDRLLPAVHVSITLREERAGAYAPLTANGTILINRYLASCYVIEEH 356

Db 426 LTKRYLVEILD--IKRVKRTGIYAPMTSQGHLLVKNKIHTSCHSEVDHH 471

# RESULT 15

719563

hypothetical protein C29F3.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T19563; T23034

R:Matthews, L.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19142

A:Accession: T19563

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-629 <Wtl>

A:Cross-references: EMBL:Z81043; PIDN:CA802804.1; GSPDB:GN00023; CESP:C29F3.2

A:Experimental source: clone C29F3

R:White, S.

submitted to the EMBL Data Library, June 1998

A:Reference number: Z19657

A:Accession: T23034

A:Status: preliminary; translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-629 <Wtl>

A:Cross-references: EMBL:AL023813; PIDN:CA19424.1; GSPDB:GN00023; CESP:C29F3.2

A:Experimental source: clone H02K04

C:Genetics:

A:Gene: CESP:C29F3.2

A:Map position: 5

A:Introns: 23/1; 111/3; 177/3; 207/2; 287/1; 381/3; 399/3; 417/1; 476/2; 528/3; 537/2

Query Match

Best Local Similarity 27.3%; Score 140.5; DB 2; Length 629;

Matches 45; Conservative 31; Mismatches 72; Indels 17; Gaps 6;

QY 199 CPGSATVHLEGGTKLVKDLRPGDRVLAADDOGRLLYSDFLTFLDRDGA 258

Db 421 CFPDAEYVNVYEKAVKRDELEVGDMVQALHGK-ETTVSPVKYWLHRDP-EQAEFVEFL 478

QY 259 LEPRERLLTAHLF-----VAPHNSGPTPGPSALFASRYRPGQRYVVAERGGDR 311

Db 479 LENGESFTLTKHLVFATDCQONKKNLDLNPST-----TKINIGE-CFMAQPENAS 531

QY 312 RLPAVHVSITLREERAGAYAPLTANGTILINRYLASCYVIEEH 356

Db 532 KFOKVQILDIG-RVRKRTGIYAPMTSLGHLVYNQIHTSCHSEIDHH 575

Search completed: February 20, 2003, 10:13:34  
Job time : 18.3815 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:07:32 ; Search time 13.4009 seconds  
(without alignments)  
1352.533 Million cell updates/sec

Title: US-09-827-110A-13

Perfect score: 2289  
Sequence: 1 MDLLARCPVLVIASSLLVC.....GTWLDSEETMPLGNVAKSS 437

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2286	99.9	437	1	SHH_MOUSE
2	2200	96.1	437	1	SHH_RAT
3	2065	90.2	462	1	SHH_HUMAN
4	1788	78.1	425	1	SHH_CHICK
5	1542.5	67.4	432	1	SHH_CYPNY
6	1525.5	66.6	444	1	SHH_XENLA
7	1517	66.3	418	1	SHH_BRARE
8	1464	64.0	416	1	THW_BRARE
9	1310	57.2	408	1	THH_CHICK
10	1271.5	55.5	411	1	THH_HUMAN
11	1268	55.4	411	1	THH_MOUSE
12	1267	55.4	409	1	THH_XENLA
13	1228	53.6	412	1	THH_BRARE
14	1219	53.3	386	1	DHH_MOUSE
15	1217	53.2	396	1	DHH_HUMAN
16	1104.5	48.3	398	1	DHH2_XENLA
17	1098.5	48.0	396	1	DHH1_XENLA
18	894	39.1	471	1	HH_DROME
19	855.5	37.4	481	1	HH_DROHY
20	603	26.3	121	1	SHH_RASEL
21	600	26.2	121	1	SHH_CARAU
22	600	26.2	121	1	SHH_PUNTE
23	599	26.2	121	1	SHH_RASHE
24	599	26.2	121	1	SHH_RASPA
25	598	26.1	121	1	SHH_TANAL
26	596	26.0	121	1	SHH_AMCH
27	595	26.0	121	1	SHH_DANVA
28	595	26.0	121	1	SHH_DANAT
29	595	26.0	121	1	SHH_DANAT
30	595	26.0	121	1	SHH_DANER
31	595	26.0	121	1	SHH_DANKE
32	595	26.0	121	1	SHH_DANPU
33	595	26.0	121	1	SHH_DEVDE

34	595	26.0	121	1	SHH_DEVMA	013247 devario mal
35	595	26.0	121	1	SHH_DEVPA	013250 devario pat
36	591	25.8	121	1	SHH_PUNCO	013250 devario con
37	399	17.4	88	1	DHH_BRARE	P79729 brachydantio
38	286	12.9	58	1	THH_CARAU	P79693 carassius a
39	286	12.9	58	1	THH_DANAT	013240 carassius a
40	296	12.9	58	1	THH_DANKE	P79711 dantio aff.
41	296	12.9	58	1	THH_DANKE	P79719 dantio kerri
42	296	12.9	58	1	THH_DEVDE	013243 devario dev
43	296	12.9	58	1	THH_PUNTE	P79852 devario tet
44	286	12.9	58	1	THH_RASEL	P79860 raspora ele
45	291	12.7	58	1	THH_DEVMA	013215 devario ma1

## ALIGNMENTS

RESULT 1  
SHH\_MOUSE  
ID 062226;  
AC 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
GN Sonic hedgehog protein precursor (SHH) (HHG-1).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Embryo;  
RX MEDLINE=94094334; PubMed=7916661;  
RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Moller J.,  
RT McMahon J.A., McMahon A.P.,  
RT "Sonic hedgehog, a member of a family of putative signaling  
molecules, is implicated in the regulation of CNS polarity.";  
RL Cell 75:1417-1430(1993).  
RN [2]  
RP REVISION TO 122.  
RC STRAIN=C57BL/6J;  
RA McMahon A.P.;  
RN Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
RX [3]  
RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.  
RX MEDLINE=95236997; PubMed=7720571;  
RA Zhang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,  
RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;  
RT "Products, genetic linkage and limb patterning activity of a murine  
hedgehog gene.";  
RL Development 120:3339-3353(1994).  
RN [4]  
RP FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.  
RX MEDLINE=95254654; PubMed=7736596;  
RA Roellink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,  
RT Beachy P.A., Jessell T.M.;  
RT "Floor plate and motor neuron induction by different concentrations of  
the amino-terminal cleavage product of sonic hedgehog  
autoproteolysis.";  
RL Cell 81:445-455(1995).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 34-195.  
RX MEDLINE=96069744; PubMed=7477329;  
RA Hall T.M.T., Porter J.A., Beachy P.A., Leahy D.J.;  
RT "A potential catalytic site revealed by the 1.7-A crystal structure of  
the amino-terminal signalling domain of Sonic hedgehog.";  
RL Nature 378:212-216(1995).  
RN [6]  
RP FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN  
ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION  
OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSES THE  
CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER  
TARGET, THE G11 ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A

VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY SIMILARITY).

-1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).

-1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF EMBRYONIC TISSUES INCLUDING THE NOTOCHORD, VENTRAL NEURAL TUBE, FLOOR PLATE, LUNG BUD, ZONE OF POLARIZING ACTIVITY AND POSTERIOR DISTAL MESENCHYME OF LIMBS. IN THE ADULT, EXPRESSED IN LUNG AND NEURAL RETINA.

-1- DEVELOPMENTAL STAGE: FIRST DETECTABLE DURING GASTRULATION.

-1- INDUCTION: BY RETINOIC ACID.

-1- PM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY.

-1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

-----

THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION - THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X76290; CAA53922.1; -  
 PDB; 1VHH; 29-JAN-96.  
 DR MEROPS; C46.002; -  
 DR MGD; MGI:98297; Shh.  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hnt.  
 DR InterPro: IPR003586; Hedgehog\_hntc.  
 DR InterPro: IPR003587; Hedgehog\_hntn.  
 DR InterPro: IPR002203; Intein.  
 DR InterPro: IPR001657; SonicHH.  
 DR Pfam: PF01079; Hnt; 1.  
 DR Pfam: PF01085; HH\_signal; 1.  
 DR PRINTS; PR00632; SONICHHOG.  
 DR PRODOM; PD003042; HH\_signal; 1.  
 DR SMART; SM00305; Hntc; 1.  
 DR SMART; SM00306; Hntn; 1.  
 DR PROSITE; PS50817; INTEIN\_NTER; 1.  
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KW Signal; Lipoprotein; Palmitate; 3D-structure.  
 FT SIGNAL 1 24  
 FT CHAIN 25 437  
 FT CHAIN 25 198  
 FT CHAIN 199 437  
 FT SITE 198 199  
 FT SITE 244 244  
 FT SITE 268 268  
 FT ACT\_SITE 271 271  
 FT BINDING 198 198  
 FT DOMAIN 383 387  
 FT LIPID 25 25  
 FT SEQUENCE 437 AA; 47773 MW; DOB872F08E7860EF CRC64;

Query Match 99.9%; Score 2286; DB 1; Length 437;

Best Local Similarity 99.9%; Pred. No. 3e-176;  
 Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MILLARCEFLVILASSLLVCGIACGPGRGKRRHPKPLIPYKOTIPVAEKTIGAS 60  
 |||||||  
 Db 1 MILLARCEFLVILASSLLVCGIACGPGRGKRRHPKPLIPYKOTIPVAEKTIGAS 60  
 |||||||

Oy 61 GRYECKITRNSEREKELTPNPNPDIFNDEBNTGADRLMTORCKDKLNAISVNWMPG 120  
 |||||||  
 Db 61 GRYECKITRNSEREKELTPNPNPDIFNDEBNTGADRLMTORCKDKLNAISVNWMPG 120  
 |||||||

Oy 121 VRLRTEGMDDEDGHHSESLIYEGEAVDITSDRKSRYGMALRLAVAGDWTYTESKA 180  
 |||||||  
 Db 121 VRLRTEGMDDEDGHHSESLIYEGEAVDITSDRKSRYGMALRLAVAGDWTYTESKA 180  
 |||||||

Oy 181 HHCVKAKNSVAAKSGCGPGSATVHLGEGTKLVKRLRGDRVLADDDGRLIYSDFL 240  
 |||||||  
 Db 181 HHCVKAKNSVAAKSGCGPGSATVHLGEGTKLVKRLRGDRVLADDDGRLIYSDFL 240  
 |||||||

Oy 241 TFLDRDEGAKKVFYIETLEPRERLLTAHLLFVAPHNDSGPTPGSPALASRYRPGQR 300  
 |||||||  
 Db 241 TFLDRDEGAKKVFYIETLEPRERLLTAHLLFVAPHNDSGPTPGSPALASRYRPGQR 300  
 |||||||

Oy 301 VYVAERGGDRLLPAAVHSTLAREEAGATPLTAHETIILNRYLASCYAVIEHSMAN 360  
 |||||||  
 Db 301 VYVAERGGDRLLPAAVHSTLAREEAGATPLTAHETIILNRYLASCYAVIEHSMAN 360  
 |||||||

Oy 361 RAFAPFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPTAGIMYSQLYHIGTW 420  
 |||||||  
 Db 361 RAFAPFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPTAGIMYSQLYHIGTW 420  
 |||||||

Oy 421 LIDSETMHPILGMAYKSS 437  
 |||||||  
 Db 421 LIDSETMHPILGMAYKSS 437  
 |||||||

RESULT 2  
 SHH\_RAT  
 ID SHH\_RAT STANDARD; PRT; 437 AA.  
 AC Q63673;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sonic hedgehog protein precursor (SHH).  
 GN SHH OR VHH-1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Embryonic floor plate;  
 RX MEDLINE=94170375; PubMed=8124714;  
 RA Roelink H., Augsburg A., Heemskerk J., Korzh V., Norlin S.,  
 RA Ruiz I., Altada A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,  
 RA Dodd J.;  
 RT "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog  
 RT of hedgehog expressed by the notochord.";  
 RL Cell 76:761-775(1994).  
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN  
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION  
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE  
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER  
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A  
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED  
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE  
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE  
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH  
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD  
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS  
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY  
 CC SIMILARITY).

-1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE



RP VARIANTS HPE ARG-31; GLY-117; ARG-117.  
 RX MEDLINE-97051937; PubMed-8886572;  
 RA Roessler E., Belloni E., Gaudenz K., Jay P., Berta P., Scherer S.W.,  
 RA Tsui L.-C., Muenke M.;  
 RT "Mutations in the human Sonic hedgehog gene cause holoprosencephaly.";  
 RL Nat. Genet. 14:357-360(1996).  
 RN [71]  
 RP VARIANTS HPE ARG-31; GLY-117; ARG-117; GLU-224; THR-226 AND THR-383.  
 RX MEDLINE-98027056; PubMed-9302282;  
 RA Roessler E., Belloni E., Gaudenz K., Vargas F., Scherer S.W.,  
 RA Tsui L.-C., Muenke M.;  
 RT "Mutations in the C-terminal domain of Sonic hedgehog cause  
 RT holoprosencephaly.";  
 RL Hum. Mol. Genet. 6:1847-1853(1997).  
 RN [81]  
 RP VARIANTS HPE HIS-100; GLN-188 AND ASN-222.  
 RX MEDLINE-99371775; PubMed-10441331;  
 RA Odent S., Atti-Bitach T., Blayau M., Mathieu M., Aug J.,  
 RA Delgado A.L., Gall J.Y., Le Marec B., Munnich A., David V.,  
 RA Vekemans M.;  
 RT "Expression of the Sonic hedgehog (SHH) gene during early human  
 RT development and phenotypic expression of new mutations causing  
 RT holoprosencephaly.";  
 RL Hum. Mol. Genet. 8:1683-1689(1999).  
 RN [91]  
 RP VARIANTS HPE V-88; K-115; R-236; 263-R--A-269 DEL; D-290; A-424 AND  
 RP L-436.  
 RX MEDLINE-20025757; PubMed-10556296;  
 RA Nann L., Ming J.E., Bocian M., Steinhans K., Bianchi D.W.,  
 RA Die-Smulders C., Giannotti A., Imizumi K., Jones K.L., Campo M.D.,  
 RA Martin R.A., Meincke P., Pierpont M.E.M., Robin N.H., Young I.D.,  
 RA Roessler E., Muenke M.;  
 RT "The mutational spectrum of the sonic hedgehog gene in  
 RT holoprosencephaly: SHH mutations cause a significant proportion of  
 RT autosomal dominant holoprosencephaly.";  
 RL Hum. Mol. Genet. 8:2479-2488(1999).  
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN  
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION  
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE  
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER  
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A  
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED  
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE  
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE  
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BOD. THRESHOLD  
 CC FLOOR PLATE-AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD  
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS  
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN FETAL INTESTINE, LIVER, LUNG, AND  
 CC KIDNEY. NOT EXPRESSED IN ADULT TISSUES.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- DISEASE: DEFECTS IN SHH ARE THE CAUSE OF HOLOPROSENCEPHALY (HPE).  
 CC HPE IS A CLINICALLY VARIABLE AND GENETICALLY HETEROGENEOUS  
 CC MALFORMATION IN WHICH THE DEVELOPING FOREBRAIN FAILS TO CORRECTLY  
 CC SEPARATE INTO RIGHT AND LEFT HEMISPHERES. IN THE MOST SEVERE FORM  
 CC (ALOBAR HPE), THERE IS NO INTERHEMISPHERIC FISSURE, A SINGLE BRAIN  
 CC VENTRICLE IS PRESENT. HPE IS ASSOCIATED WITH SEVERAL DISTINCT  
 CC FACIES AND PHENOTYPIC VARIABILITY. IN THE MOST EXTREME CASES,  
 CC ANOPHTHALMIA OR CYCLOPIA IS EVIDENT ALONG WITH A CONGENITAL  
 CC ABSENCE OF THE MATURE NOSE. THE LESS SEVERE FORM FEATURES FACIAL

CC		DSMOPLA CHARACTERIZED BY OCULAR HYPERTELORISM, DEFECTS OF THE
CC		UPPER LIP AND/OR NOSE, AND ABSENCE OF THE OLFACTORY NERVES OR
CC		CORPUS CALLOSUM. THE MAJORITY OF HPE CASES ARE APPARENTLY
CC		SPORADIC, ALTHOUGH CLEAR EXAMPLES OF AUTOSOMAL DOMINANT (AD)
CC		INHERITANCE HAVE BEEN DESCRIBED. INTERESTINGLY, UP TO 30% OF
CC		OBLIGATE CARRIERS OF HPE GENE IN AD PEDIGREES ARE CLINICALLY
CC		UNAFFECTED.
CC	-1-	SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC	-1-	DATABASE: NAME-Atlas Genet. Cyogenet. Oncol. Haematol.;
CC		WWW= <a href="http://www.infobiogen.fr/services/chromancer/Genes/SHHD378.htm">http://www.infobiogen.fr/services/chromancer/Genes/SHHD378.htm</a> "
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC		the European Bioinformatics Institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (See <a href="http://www.isdb.ch/announce/">http://www.isdb.ch/announce/</a>
CC		or send an email to <a href="mailto:license@isdb-sib.ch">license@isdb-sib.ch</a> ).
CC		-----
DR	EMBL; L38518; AAA62179.1; --	
DR	EMBL; AC002484; AAB67604.1; --	
DR	HSSP; Q62226; IVHH.	
DR	MEROES; C46_002; --	
DR	Genew; HGNC:10848; SHH.	
DR	MIM; 600725; --	
DR	MIM; 142945; --	
DR	Interpro; IPR000320; HH_signal.	
DR	Interpro; IPR001767; Hedgehog_hnt.	
DR	Interpro; IPR003586; Hedgehog_hlnc.	
DR	Interpro; IPR003587; Hedgehog_hlncw.	
DR	Interpro; IPR002203; Intein.	
DR	Interpro; IPR001657; SonicHH.	
DR	Pfam; PF01079; Hlntc_1.	
DR	Pfam; PF01085; HH_signal; 1.	
DR	PRINTS; PR00632; SONICHHOG.	
DR	Prodom; PD003042; HH_signal; 1.	
DR	SMART; SM00305; Hlntc_1.	
DR	SMART; SM00306; Hlntcn_1.	
DR	PROSITE; PS50817; INTEIN_N_TER; 1.	
KW	Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;	
KW	Signal; Lipoprotein; Palmitate; Disease mutation; Holoprosencephaly.	
FT	SYGNL	1 23
FT	CHAIN	24 462
FT	CHAIN	24 197
FT	CHAIN	198 462
FT	CHAIN	197 198
FT	STATE	243 243
FT	STATE	267 267
FT	ACT_SITE	270 270
FT	BINDING	197 197
FT	DOMAIN	407 411
FT	LIPID	24 24
FT	VARIANT	31 31
FT	VARIANT	88 88
FT	VARIANT	100 100
FT	VARIANT	115 115
FT	VARIANT	117 117
FT	VARIANT	117 117
FT	VARIANT	188 188
FT	VARIANT	222 222
FT	VARIANT	224 224
FT		/FTID-VAR_009168.
FT		G -> R (IN HPE).
FT		/FTID-VAR_003621.
FT		D -> V (IN HPE; FAMILIAL).
FT		/FTID-VAR_009163.
FT		O -> H (IN HPE; SPORADIC).
FT		/FTID-VAR_009164.
FT		N -> K (IN HPE; FAMILIAL).
FT		/FTID-VAR_009165.
FT		W -> G (IN HPE).
FT		/FTID-VAR_003620.
FT		W -> R (IN HPE).
FT		/FTID-VAR_003621.
FT		E -> Q (IN HPE; FAMILIAL).
FT		/FTID-VAR_009166.
FT		D -> N (IN HPE; FAMILIAL).
FT		/FTID-VAR_009167.
FT		V -> E (IN HPE).
FT		/FTID-VAR_009168.

```

FT  VARIANT  226  226  A -> T (IN HBE; FAMILIAL).
FT  VARIANT  236  236  S -> R (IN HBE; FAMILIAL).
FT  VARIANT  263  269  /FTID-VAR-009170.
FT  VARIANT  263  269  MISSING (IN HBE; SPORADIC).

Query Match  90.28; Score 2065; DB 1; Length 462;
Best Local Similarity 87.28; Pred. No. 1,9e-158;
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;

OY  2  LLLARCELYILASSLLVCPGLAGPGRGFGKRRPKKLPPLAYKOPFPVNAEKTLAGS 61
D  1  MLLARCLLVVSSLLVCSGLACGPRGFGKRRPKKLPPLAYKOPFPVNAEKTLAGS 60
OY  62  RYEGKIRNSRERKELTPNVPDIFDEENTGADRLMTORCKDKLALAIYVNMOPGV 121
D  61  RYEGKIRNSRERKELTPNVPDIFDEENTGADRLMTORCKDKLALAIYVNMOPGV 120
OY  122  RLRTGEMDEGDHSESLHTEGRAVDITTSDRSKYGMILARLAVENAGDWTYESKAH 181
D  121  KLRTGEMDEGDHSESLHTEGRAVDITTSDRSKYGMILARLAVENAGDWTYESKAH 180
OY  182  IHCSVKAENSYAAKSGGCPGSAIVHLEGGSTKLVKDLRPGDRVLADDDGRLLYDFLT 241
D  181  IHCSVKAENSYAAKSGGCPGSAIVHLEGGSTKLVKDLRPGDRVLADDDGRLLYDFLT 240
OY  242  FLDRDEGAKKVFYIETLEPERELLTAAHLLEFVAPHND-----SGTPPG 286
D  241  FLDRDEGAKKVFYIETLEPERELLTAAHLLEFVAPHNDATGEPASGSGSPGGALG 300
OY  287  PSALFASRVPRGQVYVVAERGGDRLLPAVHVSVTLEEPAGVAPLTHAGTLINRYL 346
D  301  PRAFFASRVPRGQVYVVAERGGDRLLPAVHVSVTLEEPAGVAPLTHAGTLINRYL 360
OY  347  ASCVAVEESHMAHAPAPFLAHLAALAPARTD-----GGGGGSIP-AAOSA 395
D  361  ASCVAVEESHMAHAPAPFLAHLAALAPARTDGGSGGGDRGGGGGRVALTPAGA 420
OY  396  TEARGAEPAGIHMYSQLYHIGTWLIDSETMRLPAGNAVSS 437
D  421  ADARGAGATAGIHMYSQLYIGTWLIDSETMRLPAGNAVSS 462

RESULT 4
ID  SHH_CHICK  STANDARD;  PRT;  425  AA.
AC  091035;
DT  15-JUL-1999 (Rel. 38, Created)
DT  15-JUL-1999 (Rel. 38, Last sequence update)
DT  15-OCT-2001 (Rel. 40, Last annotation update)
DE  Sonic hedgehog protein precursor (SHH).
GN  SHH.
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Limb bud;
RX  MEDLINE=9409433; PubMed=8269518;
RA  Riddle R.D., Johnson R.L., Laufer E., Tabin C.;
RT  "Sonic hedgehog mediates the polarizing activity of the ZPA.";
RL  Cell 75:1401-1416(1993).

RP  FUNCTION. AND AUTOPROTEOLYTIC CLEAVAGE.
RA  MEDLINE=95254654; PubMed=7735596;
RA  Koellink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,
RA  Beachy P.A., Jessell T.M.;
RT  "Floor plate and motor neuron induction by different concentrations of
RT  the amino-terminal cleavage product of sonic hedgehog
RT  autoproteolysis.";
RL  Cell 81:445-455(1995).

```

```

CC  -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN
CC  ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION
CC  OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE
CC  CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER
CC  TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A
CC  VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED
CC  BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE
CC  AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE
CC  ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH
CC  FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD
CC  CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS
CC  5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION.
CC  -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC  CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC  TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC  -1- TISSUE SPECIFICITY: EXPRESSED IN THE POSTERIOR LIMB BUD
CC  MESENCHYME, THE HENSEN'S NODE, THE NOTOCHORD, AND THE FLOOR PLATE
CC  OF THE NEURAL TUBE.
CC  -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT STAGE 17 DURING THE
CC  INITIATION OF LIMB BUD FORMATION. FROM THAT POINT ONWARDS, THE
CC  EXPRESSION PATTERN EXACTLY MATCHES THE LOCATION OF THE ZONE OF
CC  POLARIZING ACTIVITY (ZPA).
CC  -1- INDUCTION: BY RETINOID ACID.
CC  -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC  AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC  THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC  A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC  TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC  TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC  OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC  ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC  C-PRODUCT HAS NO SIGNALING ACTIVITY.
CC  -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: L28099; AAA72428.1; -.
DR  HSSP: O62226; LVH.
DR  MEROPS: C46.001; -.
DR  InterPro: IPR000320; HH_signal.
DR  InterPro: IPR001767; Hedgehog_hnt.
DR  InterPro: IPR003586; Hedgehog_hntC.
DR  InterPro: IPR003587; Hedgehog_hntN.
DR  InterPro: IPR002203; Intein.
DR  InterPro: IPR001657; SonICHH.
DR  Pfam: PF01079; Hnt; 1.
DR  Pfam: PF01085; HH_signal; 1.
DR  PRINTS: PR00632; SONICHOG.
DR  ProDom: PD003042; HH_signal; 1.
DR  SMART: SM00305; HntC; 1.
DR  SMART: SM00306; HntN; 1.
DR  PROSITE: PSS0817; INTEIN_N_TER; 1.
KW  Developmental Protein; Autocatalytic cleavage; Hydrolase; Protease;
KW  Signal; Lipoprotein; Palmitate.
FT  SIGNAL 1 26
FT  CHAIN 27 425
FT  CHAIN 27 200
FT  CHAIN 201 425
FT  SITE 200 201
FT  SITE 246 246
FT  SITE 270 270
FT  ACT_SITE 273 273
FT  BINDING 200 200

```

FT DOMAIN 390 393 POLY-THR.  
 FT LIPID 27 27 PALMITATE (BY SIMILARITY).  
 SQ SEQUENCE 425 AA: 46474 MW: DA9627443DA0173 CRC64:

Query Match 78.1%; Score 1788; DB 1; Length 425;  
 Best Local Similarity 80.2%; Pred. No. 3e-136;  
 Matches 353; Conservative 22; Mismatches 43; Indels 22; Gaps 6;

QY 2 LLLARCFVLVIASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTGLGASG 61  
 DB 4 MLLRLRLVLFICALVSSGLTGPGRGIGKRRHPKLTPLAYKQFIPNVAEKTGLGASG 63  
 QY 62 REGGRTIRSEFKELTPYNPNDIIFKDEBNGADRLMTQKCKDLNLAISVMNQMGV 121  
 DB 64 REGGRTIRSEFKELTPYNPNDIIFKDEBNGADRLMTQKCKDLNLAISVMNQMGV 123  
 QY 122 RLRTVEGDEDEHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDVYVESKAH 181  
 DB 124 KLRTVEGDEDEHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDVYVESKAH 183  
 QY 182 IHCYKAENSVAAGSGGCPGSAVTHLEGGTKLVKDLRPGDRLAADGRLVSDFLT 241  
 DB 184 IHCYKAENSVAAGSGGCPGSAVTHLEGGTKLVKDLRPGDRLAADGRLVSDFLT 243  
 QY 242 FLDRBEGAKKVVYIETLEPRERLLITAAHLFVAP-HNDS--GPTGPGSLFASRVAP 297  
 DB 244 FLDRDSSSKRLFYETROPARALLTAHLFVAPHNQSEATGSGS-QALFASVWAP 302  
 QY 298 GORVYVAERGGDRLLPAVHVTLRFEAGAVAPLTAHGTLINRYLASGYAIEEHS 357  
 DB 303 GORVYVAGE--GGQQLPVSIVHSVLSREPAQVAPLTAHGTLINRYLASGYAIEEHS 360  
 QY 358 WAHRAFAFRLAHLAALAPARTDGGGSGIPAAQSATEAGAPTAGIHVYSQLYHI 417  
 DB 361 WAHMAFAFRLAQLLAALCP-----DGAIPFA-----ATTGTSIHVYSLYHI 405  
 QY 418 GTWLDDSEMHPLGMAVKSS 437  
 DB 406 GSWVLDDGALHPLGMVAPAS 425

RESULT 5  
 SHH\_CYNPY STANDARD: PRT: 432 AA.  
 AC 090385;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sonic hedgehog protein precursor (SHH).  
 GN SHH.  
 OS Cynops pyrrhogaster (Japanese common newt).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Caudata; Salamandroides; Salamandridae; Cynops.  
 OX NCBI\_TaxID=8330;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RG MEDLINE=96136334; PubMed=8573168;  
 RX Takabatake T., Takahashi T.C., Inoue K., Ogawa M., Takeshima K.;  
 RA "Activation of two Cynops genes, fork head and sonic hedgehog, in  
 RT animal cap explants.";  
 RL Blochem. Biophys. Res. Commun. 218:395-401(1996).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED  
 CC IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND  
 CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND  
 CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH  
 CC FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE  
 CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC  
 CC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE

CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- INDUCTION: ACTIVATED BY ACTIVIN, BASIC FIBROBLAST GROWTH FACTOR  
 CC (BFGF) AND FORK HEAD.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: D63339; BA09657.1; -

CC HSSP: Q62226; IYH.

CC MEROPS: C46.002; -

CC InterPro: IPR000320; HH\_signal.

CC InterPro: IPR001767; Hedgehog\_hntc.

CC InterPro: IPR003586; Hedgehog\_hntc.

CC InterPro: IPR003587; Hedgehog\_hntc.

CC InterPro: IPR002203; Inteln.

CC InterPro: IPR001657; SonicHH.

CC Pfam: PF01079; Hint. 1.

CC PRINTS: PR00185; HH\_signal. 1.

CC PRODOM: PD003042; HH\_signal. 1.

CC SMART: SM00305; Hntc. 1.

CC SMART: SM00306; Hntc. 1.

CC PROSITE: PS00817; INTEN\_N\_TER. 1.

CC Signal: Lipoprotein; Palmitate.

CC SIGNAL 1 26

CC CHAIN 27 432

CC CHAIN 27 200

CC CHAIN 201 432

CC SITE 200 201

CC SITE 268 268

CC ACT\_SITE 271 271

CC BINDING 200 200 CHOLESTEROL (BY SIMILARITY).  
 CC LIPID 27 27 PALMITATE (BY SIMILARITY).  
 SQ SEQUENCE 432 AA: 47847 MW: BA55C7E746C8B5A8 CRC64:

Query Match 67.4%; Score 1542.5; DB 1; Length 432;  
 Best Local Similarity 70.2%; Pred. No. 1.7e-116;  
 Matches 313; Conservative 32; Mismatches 74; Indels 27; Gaps 7;

QY 2 LLLARCFVLVIASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTGLGASG 61  
 DB 4 MLLRLRLVLFICALVSSGLTGPGRGIGKRRHPKLTPLAYKQFIPNVAEKTGLGASG 63  
 QY 62 REGGRTIRSEFKELTPYNPNDIIFKDEBNGADRLMTQKCKDLNLAISVMNQMGV 121  
 DB 64 REGGRTIRSEFKELTPYNPNDIIFKDEBNGADRLMTQKCKDLNLAISVMNQMGV 123  
 QY 122 RLRTVEGDEDEHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDVYVESKAH 181  
 DB 124 KLRTVEGDEDEHSESLHYEGRAVDITTSDRSKYGMRLAVEAGFDVYVESKAH 183  
 QY 182 IHCYKAENSVAAGSGGCPGSAVTHLEGGTKLVKDLRPGDRLAADGRLVSDFLT 241  
 DB 184 IHCYKAENSVAAGSGGCPGSAVTHLEGGTKLVKDLRPGDRLAADGRLVSDFLT 243



Oy	242	FLNDEGAKKVFYIEFLREKRELLTAHLFLVFA----	PHNDSG-----	PTPEPSALF	291
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :		
Db	244	FMDKEEYVRYKVFYIEFL--SREKRYRLTAHLFLV	FGAHPHNSGGDGFNSVGSAGFRNMF		301
Oy	292	ASRYRPGQRYVVAENRGCDRRLLPAAVHVSUTL	NEEDGAGYADLTGAHGTILINRYLASCYA	351	
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :		
Db	302	ASSVRAGHRYLVLTVDREG--RGLREAFVIERYLL	EAATGAYAPVTAHGTIVDRVLASCYA	358	
Oy	332	VIEEHSNHNRAFPFRLAHLAALPARTDGGGGS	IPAAOSATDANGAEPTACIMYWS	411	
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :		
Db	359	VIEEHSNHNRAFPFRLAHLAALPARTDGGGGS	IPAAOSATDANGAEPTACIMYWS	411	
Oy	412	OLLYHGTWLLDSETHPLDLMMAVKS	437		
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :		
Db	407	EILRIGTWIQLDIEDTHPLDLMMAKSS	432		
		: : : : : : : : : : : : : : : : : :	: : : : : : : : : : : : : : : : : :		
RESULT 6					
SHH_XENLA		STANDARD:	PRT:	444 AA.	
ID	092000:	091894:			
AC	15-JUL-1999	(Rel. 38, Created)			
DT	15-JUL-1999	(Rel. 38, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	SHH hedgehog protein precursor (X-SHH) (VHH-1).				
GN	SHH.				
OS	Xenopus laevis (African clawed frog).				
OC	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;				
OC	Amphibia: Batrachia: Anura: Mesobatrachia: Pipidoae; Pipidae;				
OC	Xenopodidae; Xenopus.				
OX	NCBI_Taxid=8355;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Intestine;				
RX	MEDLINE=95357169; PubMed=7630736;				
RA	Stolow M.A., Shi Y.-B.;				
RT	"Xenopus sonic hedgehog as a potential morphogen during embryogenesis				
RT	and thyroid hormone-dependent metamorphosis.";				
RL	Nucleic Acids Res. 23:2555-2562(1995).				
	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Embryo;				
RX	MEDLINE=95401852; PubMed=7671800;				
RA	Ecker S.C., McGrew L.L., Lal C.-J., Lee J.J., von Kessler D.P.,				
RT	Moon R.T., Beachy P.A.;				
RT	"Distinct expression and shared activities of members of the hedgehog				
RT	gene family of Xenopus laevis.";				
RL	Development 121:2337-2347(1995).				
	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Notochord;				
RX	MEDLINE=96028338; PubMed=7551564;				
RA	Ruitz I Altaba A., Jessell T.M., Roelink H.;				
RT	"Restrictions to floor plate induction by hedgehog and winged-helix				
RT	genes in the neural tube of frog embryos.";				
RL	Mol. Cell. Neurosci. 6:106-121(1995).				
CC	-1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF				
CC	PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED				
CC	IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND				
CC	VENTRAL SOMITE DIFFERENTIATION. INDUCES ECOTOPIC CEMENT GLAND				
CC	FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH				
CC	FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE				
CC	TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC				
CC	REPRESSSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY				
CC	STIMILIARITY).				
CC	-1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE				
CC	CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE				
CC	CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM				
CC	TO MEDIANE SIGNALING TO OTHER CELLS (BY STIMILIARITY).				
CC	-1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN NOTOCHORD AND NEURAL				
CC	FLOOR PLATE DURING EMBRYOGENESIS. IN TADPOLE, HIGH EXPRESSION IS				
CC	OBSERVED IN PANCREAS/STOMACH MODERATE EXPRESSION IN TAIL, AND LOW				

CC		EXPRESSIN IN INTESTINE, BRAIN, AND HIND LIMB.
CC	-1-	DEVELOPMENTAL STAGE: FIRST DETECTED AT THE NEURULA (STAGES 16-17).
CC		FIRST PEAK OF EXPRESSION AROUND TADPOLE HATCHING (STAGES 33-40).
CC		HIGH EXPRESSION OBSERVED IN INTESTINE AT THE CLIMAX OF MORPHOGENESIS (STAGES 60-62) WHEN INTESTINE EPITHELIAL UNDERGOES MORPHOGENESIS.
CC	-1-	INDUCTION: BY THYROID HORMONE.
CC	-1-	PMM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC		SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/isb-sib.ch">http://www.isb-sib.ch/announce/isb-sib.ch</a> ).
CC		or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .
CC		-----
DR	EMBL; L39213;	AAC842227.1; -
DR	EMBL; U26314;	AAA85162.1; -
DR	EMBL; L35248;	AAA4981.1; -
DR	HSSP; O62226;	IYHH.
DR	MEROPS; C46.002;	--
DR	InterPro; IPR000320;	HL_signal.
DR	InterPro; IPR001767;	Hedgehog_hlnt.
DR	InterPro; IPR003586;	Hedgehog_hlntc.
DR	InterPro; IPR003587;	Hedgehog_hlntn.
DR	InterPro; IPR002203;	Inteln.
DR	InterPro; IPR001657;	SonicHH.
DR	Pfam; PF01079;	Hlnt; 1.
DR	Pfam; PF01085;	HL_signal; 1.
DR	PRINTS; PR00632;	SONICHOG.
DR	ProDom; PD003042;	HL_signal; 1.
DR	SMART; SMO0305;	HlntC; 1.
DR	SMART; SMO0306;	Hlntcn; 1.
DR	PROSITE; PS50817;	INTER_N_TER; 1.
KW	Developmental protein; Autocatalytic cleavage; Hydrolase; Protease; signal; Lipoprotein; Palmitate; Repeat.	
FT	SIGNAL	1 24 POTENTIAL.
FT	CHAIN	25 444 SONIC HEDGEHOG PROTEIN.
FT	CHAIN	25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.
FT	CHAIN	199 444 SONIC HEDGEHOG PROTEIN C-PRODUCT.
FT	SITE	198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT	SITE	266 INVOLVED IN AUTO-CLEAVAGE (BY SIMILARITY).
FT	ACT_SITE	269 ESSENTIAL FOR AUTO-CLEAVAGE (BY SIMILARITY).
FT	DOMAIN	386 409 3 X 8 AA TANDEM REPEATS OF Q-V-D-L-Q-S-H-H.
FT	REPEAT	386 393 1.
FT	REPEAT	394 401 2.
FT	REPEAT	403 409 3.
FT	BINDING	198 198 CHOLESTEROL (BY SIMILARITY).
FT	LIPID	25 25 PALMITATE (BY SIMILARITY).
FT	CONFLICT	5 9 TOSLI -> NSNLDM (IN REF. 3).
FT	CONFLICT	302 319 DPKMTTLKAAVEVVDLE -> ESQDHDLEGGRWRLLILR (IN REF. 3).
FT	CONFLICT	N -> S (IN REF. 3).
SO	SEQUENCE	444 AA; 49453 MW; 73ABE4932FEA2EEF2 CRC64;
Query Match		Score 1525.5; DB 1; Length 444;
Best Local Similarity	67.48;	Pred. No. 4e-115;
Matches 304;	Conservative 38;	Mismatches 88; Indels 21; Gaps 6;

```

DB      1  MVAATGSLLLSPICLTPPCACGPGRGICRRPRPKLTPLAYKOFIPNVAEKLIGAS 60
OY      61  GYEGKITRNSERKEITNPNYNDIIFKDEMTGADRLTORCKDKLNLAISVMQWQ 120
DB      61  GYEGKITRNSERKEITNPNYNDIIFKDEMTGADRLTORCKDKLNLAISVMQWQ 120
OY      121  VALRTVEGDEGDHSEESLHYEGRAVDITTSDRSKYGMRLARLAVEGFMVVEESKA 180
DB      121  VALRTVEGDEGDHSEESLHYEGRAVDITTSDRSKYGMRLARLAVEGFMVVEESKA 180
OY      181  HHCSYKAANSVAASGCGFPGSATYHLQCGTKLYKDLRPGDRVLAADOGRLYSDFL 240
DB      181  HHCSYKAANSVAASGCGFPGSATYHLQCGTKLYKDLRPGDRVLAADOGRLYSDFL 240
OY      241  TELDEGAKKFEYVETLEPERERLLTAHLLEFVAPHNDSGPTPSALSRVPRGOR 300
DB      241  MLDGDRDVKKLFYET--SQRKITLTAHLLEFVAGTQVNG--TRSKVFPASNIDPGDL 297
OY      301  VYVVAERGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILINRLVASCYAVIEHSMW 360
DB      301  VYVVAERGDRRLPAAVHSVTLREBEAGAVAPLTAHGTILINRLVASCYAVIEHSMW 360
OY      361  RAFPRLAHALLAALAPARTDGGCGGSTRPAQ-----SATFARGAPRTAG 406
DB      361  RAFPRLAHALLAALAPARTDGGCGGSTRPAQ-----SATFARGAPRTAG 406
OY      407  IHMYSQLYHIGTWLIDSETMHPGLMAVYKS 437
DB      414  IHMYSQLYHIGTWLIDSETMHPGLMAVYKS 444

```

```

RESULT 7
SHH_BRARE
ID AC 092008: 013170: 013171:
AC 15-JUL-1998 (Rel. 38, Created)
DT 15-JUL-1998 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sonic hedgehog protein precursor (SHH) (VNH-1).
GN SHH OR VNH1.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=embryo;
RA MEDLINE=94170375; PubMed=8124714;
RA Roelink H., Augsburg A., Heemskerk J., Korzh V., Nollin S.,
RA Ruiz I Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
RA Dodd J.;
RT "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog
RT of hedgehog expressed by the notochord.";
RL Cell 76:761-775(1994).
RN [2]
RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RA MEDLINE=96014264; PubMed=7583153;
RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P., Porter J.A.,
RA Moon R.T., Beachy P.A.;
RT "Patterning activities of vertebrate hedgehog proteins in the
RT developing eye and brain.";
RL Curr. Biol. 5:944-955(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=96083328; PubMed=7579523;
RA Fietz M.J., Concordet J.-P., Barbosa R., Johnson R., Krauss S.,
RA McMahon A.P., Tabin C., Ingham P.W.;
RT "The hedgehog gene family in Drosophila and vertebrate development.";
RL Development Suppl. 43-51(1994).
RN [4]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=99225458; PubMed=10207136;
RA Muller F., Chang B., Albert S., Fischer N., Tora L., Strahle U.;
RT "Intronic enhancers control expression of zebrafish sonic hedgehog in
RT floor plate and notochord.";
RL Development 126:2103-2116(1999).
RN [5]
RP SEQUENCE OF 30-92 AND 113-170 FROM N.A.
RC TISSUE=Muscle;
RA MEDLINE=97075114; PubMed=8917540;
RA Zardoya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
RT closely related to the zebrafish.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
CC NOTOCHORD THAT INDUCES SOMITE PATTERNING. DORSO-VENTRAL PATTERNING
CC OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
CC FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
CC PTC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
CC TUBE AND BRAIN. ALSO FOUND IN THE NOTOCHORD AND IN DEVELOPING FIN
CC BUD. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN DOMAINS THAT
CC INCLUDE A DISCRETE REGION IN THE FLOOR OF THE Diencephalon.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE INNER CELL LAYER OF
CC THE EMBRYONIC SHIELD DURING GASTRULATION. BY 9.5 HRS OF
CC DEVELOPMENT, EXPRESSED IN A CONTINUOUS BAND THAT EXTENDS FROM THE
CC TAIL TO THE HEAD. THE ANTERIOR BOUNDARY OF EXPRESSION BEING
CC POSITIONED IN THE CENTER OF THE ANIMAL POLE ANTERIOR TO THE
CC PRESUMPTIVE MIDBRAIN.
CC -1- PTC: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L27585; AAA20998.1;
DR EMBL: U30711; AAC59742.1;
DR EMBL: Z35669; CA84738.1;
DR EMBL: AF124382; AADA7913.1;
DR EMBL: U51551; AAB38575.1;
DR EMBL: U51570; AAB38593.1;
DR HSSP: Q62226; IVNH.
DR MEROPS: C46.002;
DR ZFIN: ZDB-GENE-980526-166; shh.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001767; Hedgehog_hint.
DR InterPro: IPR003586; Hedgehog_hintC.
DR InterPro: IPR003587; Hedgehog_hintN.
DR InterPro: IPR002203; Intein.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01079; Hint; 1.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHHOG.

```



DR PRODOM: PD003042; HH\_signal; 1.  
 DR SMART; SM00305; Hintc; 1.  
 DR SMART; SM00306; Hintc; 1.  
 DR PROSITE; PSS0817; INTEIN\_N\_TER; 1.  
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KM Signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 418 SONIC HEDGEHOG PROTEIN.  
 FT CHAIN 24 197 SONIC HEDGEHOG PROTEIN N-PRODUCT.  
 FT CHAIN 198 418 SONIC HEDGEHOG PROTEIN C-PRODUCT.  
 FT SITE 197 198 CLEAVAGE (AUTO-).  
 FT SITE 243 243 INVOLVED IN CHOLESTEROL TRANSFER (BY  
 FT SITE SIMILARITY).  
 FT SITE 267 267 INVOLVED IN AUTO-CLEAVAGE (BY  
 FT ACT\_SITE 270 270 ESSENTIAL FOR AUTO-CLEAVAGE (BY  
 FT BINDING 197 197 CHOLESTEROL (BY SIMILARITY).  
 FT LIPID 24 24 PALMITATE (BY SIMILARITY).  
 SQ SEQUENCE 418 AA; 46402 MW; CF000AFPEF2E5795 CRC64;  
 Query Match 66.3%; Score 1517; DB 1; Length 418;  
 Best Local Similarity 67.7%; Pred. No. 1.8e-114; Indels 18; Gaps 3;  
 Matches 294; Conservative 47; Mismatches 75; Indels 18; Gaps 3;  
 QY 4 LLARCPVLVLAASSLLVPGIACGPRGFRGRRHKKLPLAYKQFIPNVAEKTGASGRY 63  
 DB 3 LRTVLVLSLTLSLVVSGIACGPRGRRHKKLPLAYKQFIPNVAEKTGASGRY 62  
 QY 64 EGGTRNSERKELTPYNDIIFKDEENGADRLMTORCKDKNALALISVMNMPGVR 123  
 DB 63 EGGTRNSERKELTPYNDIIFKDEENGADRLMTORCKDKNALALISVMNMPGVR 122  
 QY 124 RTEGMDEDGHHSESLHYEGRAVDITTSDBRSKYGLARLAVEAGPDWYVESKAHII 183  
 DB 123 RVTGCMEDGHHSESLHYEGRAVDITTSDBRSKYGLARLAVEAGPDWYVESKAHII 182  
 QY 184 CSYKENSVAKSGCGEPGSAIVHLEQGTLYKLVLDLRPGDVLAAADOGRLLYSDFTFL 243  
 DB 183 CSYKENSVAKSGCGEPGSAIVHLEQGTLYKLVLDLRPGDVLAAADOGRLLYSDFTFL 242  
 QY 244 DRDGAARVPIVITELPERRELLITAAHLFVAPANDSGPRPSALFASVVRGQRYV 303  
 DB 243 DRDGTTRVPIVITELPERRELLITAAHLFVAPANDSGPRPSALFASVVRGQRYV 301  
 QY 304 VAEGGDRRLPLAAVHSVTLREEGAGVAPLTAHGTILINVLASCYAVIEHSHARAF 363  
 DB 302 VDSGQLKSVLVQRIYI---EQRGSAFAPTAHGTIVVDRIILASCYAVIEDGLAHAF 357  
 QY 364 AFRILAHALLAALAPARTDGGGSGIPASASATEARGAEPAGIHWYSQLLYHIGTWLID 423  
 DB 358 AFRILYVVSFLPP-----QNSSRSNATLQEGGVHWSRLLYQGTWILD 404  
 QY 424 SETMHPLGMAYKSS 437  
 DB 405 SNMLPLGMSVNSS 418  
 RESULT 8  
 TWMH\_BRARE STANDARD; PRT; 416 AA.  
 AC Q90419;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Tiggy-winkle hedgehog protein precursor (TWMH).  
 GN TWMH.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_Taxid=7955;  
 RN [1]

RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.  
 RC TISSUE-Embryo;  
 RX MEDLINE-96014264; PubMed-758133;  
 RA Ecker S.C., Ungar A.R., Greenstein P., von Kessler D.P.,  
 RT Porter J.A., Moon R.T., Beachy P.A.; hedgehog proteins in the  
 RT developing eye and brain.";  
 RL Curr. Biol. 5:944-955(1995).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT. INVOLVED IN DORSO-VENTRAL  
 CC PATTERNING OF THE BRAIN AND IN EARLY PATTERNING OF THE DEVELOPING  
 CC EYES.  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL  
 CC TUBE AND BRAIN. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN  
 CC DOMAINS THAT INCLUDE A DISCRETE REGION IN THE FLOOR OF THE  
 CC DIENCEPHALON. NOT DETECTED IN THE NOTOCHORD OR DEVELOPING FIN BUD.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U30710; AAC59741.1; -.  
 CC HSSP: Q62226; 1VHH.  
 DR MEROPS; C46.001; -.  
 DR ZFIN; ZDB-GENE-980526-41; twmh.  
 DR InterPro; IPR000320; HH\_signal.  
 DR InterPro; IPR001767; Hedgehog\_hintc.  
 DR InterPro; IPR003586; Hedgehog\_hintc.  
 DR InterPro; IPR003587; Hedgehog\_hintc.  
 DR InterPro; IPR001657; Sonichn.  
 DR Pfam; PF01079; Hintc; 1.  
 DR Pfam; PF01085; HH\_signal; 1.  
 DR PRINTS; PR00632; SONICHOG.  
 DR PRODOM; PD003042; HH\_signal; 1.  
 DR SMART; SM00305; Hintc; 1.  
 DR SMART; SM00306; Hintc; 1.  
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KM Signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 26  
 FT CHAIN 27 416 POTENTIAL.  
 FT CHAIN 201 416 TIGGY-WINKLE HEDGEHOG PROTEIN N-PRODUCT.  
 FT SITE 200 201 TIGGY-WINKLE HEDGEHOG PROTEIN C-PRODUCT.  
 FT SITE 270 270 CLEAVAGE (AUTO-).  
 FT SITE INVOLVED IN AUTO-CLEAVAGE (BY  
 FT SITE SIMILARITY).  
 FT ACT\_SITE 273 273 ESSENTIAL FOR AUTO-CLEAVAGE (BY  
 FT BINDING 200 200 CHOLESTEROL (BY SIMILARITY).  
 FT LIPID 27 27 PALMITATE (BY SIMILARITY).  
 SQ SEQUENCE 416 AA; 46576 MW; 61EC2218309CFE59 CRC64;  
 Query Match 64.0%; Score 1464; DB 1; Length 416;  
 Best Local Similarity 66.7%; Pred. No. 3.2e-110; Indels 36; Gaps 6;  
 Matches 287; Conservative 43; Mismatches 64; Indels 36; Gaps 6;



```

OY 359 AHRFAFRLAHLAALAPRTDGGGGSIPAAQSATEARGAPTAGIHMYSQLYHIC 418
DB 354 AQMAFWRLRYHSLD-----GG--PGVG-----DGVHYSGLYRLG 389
OY 419 TWLIDSETHPLG 431
DB 390 RMLLPDPSPHRLG 402

RESULT 10
ID IHH_HUMAN STANDARD; PRT; 411 AA.
AC Q14623; O43322;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Indian hedgehog protein precursor (IHH) (HHG-2).
GN IHH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tate G., Kishimoto K., Mitsuura T.;
RT "Expression of Sonic hedgehog and its receptor Patched/smoothed in
RT human cancer cell lines and embryonic organs.";
RL J. Blochem. Mol. Biol. Biophys. 4:27-34(2000).
RN [2]
RP SEQUENCE OF 100-411 FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=96070431; PubMed=7590746;
RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,
RA Gattler J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seldman C.E.,
RA Jankins N.A., Seldman J.G., McMahon A.P., Tabin C.;
RT "Cloning, expression, and chromosomal location of SHH and IHH: two
RT human homologues of the Drosophila segment polarity gene hedgehog.";
RL Genomics 28:44-51(1995).
RN [3]
RP SEQUENCE OF 124-172 FROM N.A.
RX MEDLINE=95236997; PubMed=7720571;
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandi B.K.,
RA Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;
RT "Products, genetic linkage and limb patterning activity of a murine
RT hedgehog gene.";
RL Development 120:3339-3353(1994).
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF
CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)
CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO
CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN
CC ENDOCHONDRAL OSSIFICATION. MAY REGULATE THE BALANCE BETWEEN GROWTH
CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION
CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC LUNG, AND IN ADULT
CC KIDNEY AND LIVER.
CC -1- PFM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AB018076; BAA33523.2; -
CC DR EMBL; AB010092; BAA33523.2; JOINED.
CC DR EMBL; AB018075; BAA33523.2; JOINED.
CC DR EMBL; L38517; AAA62178.1; -
CC DR HSSP; 062226; 1VHH.
CC DR MEROPS; C46.003; -
CC DR Genew; HGNC:5956; IHH.
CC DR MIM; 600726; -
CC DR InterPro; IPR000320; HH_signal.
CC DR InterPro; IPR001767; Hedgehog_hntc.
CC DR InterPro; IPR003586; Hedgehog_hntc.
CC DR InterPro; IPR003587; Hedgehog_hntc.
CC DR InterPro; IPR002203; Intein.
CC DR Pfam; PF01079; Hntc; 1.
CC DR Pfam; PF01085; HH_signal; 1.
CC DR SMART; SM00305; Hntc; 1.
CC DR SMART; SM00306; Hntc; 1.
CC DR PROSITE; PSS0817; INTEIN_N_TER; 1.
CC DR Developmental Protein; Autocatalytic cleavage; Hydrolase; Protease;
CC Signal; Lipoprotein; Palmitate.
CC FT SIGNAL 1 27
CC FT CHAIN 28 411
CC FT CHAIN 28 202
CC FT CHAIN 203 411
CC FT SITE 202 203
CC FT SITE 248 248
CC FT SITE 272 272
CC FT ACT_SITE 275 275
CC FT BINDING 202 202
CC FT LIPID 28 28
CC FT CONFLICT 100 100
CC FT CONFLICT 246 246
CC FT CONFLICT 309 309
CC FT SEQUENCE 411 AA; 4518 MW; 14C904A36C344DBA CRC64;
CC -----
CC Query Match 55.58; Score 1271.5; DB 1; Length 411;
CC Best Local Similarity 60.64; Pred. No. 9.4e-95;
CC Matches 260; Conservative 40; Mismatches 92; Indels 37; Gaps 8;
OY 8 CELVILASSLVCPL-ACGPRGFG-KRRHPKLTPLAYKQFTPNVAEKTLAGSGRYEG 65
DB 13 CLVLL--LLVVPAAAGCGGRVYGSRRRPRLVPLAYKQFSPNPEKTLAGSGRYEG 69
OY 66 KITNSRPEKLTTPNYNDITFKDEMTGADRLTORCKOKNLALSYNOWNPGVRLV 125
DB 70 KITNSRPEKLTTPNYNDITFKDEMTGADRLTORCKOKNLALSYNOWNPGVRLV 129
OY 126 TEGWDEGHSESLHAEGRVADITTSDDRSKYGMLARLAVEGPFMYYESAHHICS 185
DB 130 TEGWDEGHSESLHAEGRVADITTSDDRSKYGMLARLAVEGPFMYYESAHHICS 189
OY 186 VKAENSYAASGCGPESATVHLEGGSTKLKLDLPGRDVLADOGRLYSDFLEFLDR 245
DB 190 VKSHSAAATGCGFPAGAQVRLSCGARVALSAVRPDRVLAMGEDSPFSFVLLDDR 249
OY 246 DEGAKKFEVETLEPERILLTAHLLFVAHNDSPTEGFSALFASRRPGRGVYVA 305
DB 250 EPHRLRAFVYETDDPPRLALTPAHLLFTA-DHTEPARAFRTFASHVQPGYVLVAG 308
OY 306 ERGDRRLPAAYSVTLREEGAGAVAPLTANGTILINRYLASCYAVIEESHNAHAFAP 365
DB 309 APG---LQPARVAAS-TYHALGAVAPRLTKHGLVVEDVVASCFAAVADHHLAQLAFWP 363
OY 366 FRLAHLAALAPARTDGGGGSIPAAQSATEARGA-EPNAGIHMYSQLYHIC 424
DB 366 FRLAHLAALAPARTDGGGGSIPAAQSATEARGA-EPNAGIHMYSQLYHIC 424

```



Db 252 NLRAPVETQDPRLATPLAHLFLA-DNTERPAHFRATASHVQGVVSGVP 310  
 Qy 308 GGDRLPLPAVHVTLEEEAGAVAPLTANGTILINVLASCVAVIEHSMHARAF 367  
 Db 311 LQPARVAAS-THVALGSYAPLTRHGTLLVEDVASCFAVAADHHAOLAFLPRL 365  
 Qy 368 LAHALLAALAPARTDGGGGSIPAQSAATEARGA-EPTAGIHWYSOLLVHIGTULDS 426  
 Db 366 LFPSTL-----AMGSWTPSEGVHWPQMLYRLGRLLLEST 400  
 Qy 427 MHPILGMA 433  
 Db 401 FHPILGMS 407

RESULT 12  
 IHH\_XENLA STANDARD; PRT; 409 AA.  
 ID IHH\_XENLA  
 AC Q91612;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Indian hedgehog protein precursor (IHH) (Banded hedgehog protein) (X-BHH).  
 GN IHH OR BHH.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;  
 OC Xenopodidae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=95401852; Pubmed=7671800;  
 RA Ekker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P., Moon R.T., Beachy P.A.;  
 RT "Distinct expression and shared activities of members of the hedgehog gene family of Xenopus laevis."  
 RL Development 121:2337-2347(1995).  
 CC -1- FUNCTION: SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF ANTERODORSAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC CEMENT GLAND FORMATION IN EMBRYOS.  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDiate SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- DEVELOPMENTAL STAGE: DETECTABLE WITHIN THE EARLY GASTRULA. AT STAGE 14 NEURULA, HIGH EXPRESSION IN NEUROECTODERM. EXPRESSED THROUGHOUT THE NEURAL PLATE AND SUBSEQUENTLY IN BOTH THE NERVOUS SYSTEM AND IN THE DERMATOME OF SOMITES.  
 CC -1- INDUCTION: BY ACTIVIN.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U26404; AAA85165.1; -  
 DR HSP; Q62226; LVH.

DR MEROPS; C46.0PW; -  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hint.  
 DR InterPro: IPR003586; Hedgehog\_hintc.  
 DR InterPro: IPR003587; Hedgehog\_hintn.  
 DR InterPro: IPR001657; SonICHR.  
 DR Pfam: PF01079; HH\_signal.  
 DR Pfam: PF01085; HH\_signal.  
 DR PRINTS: PR006342; SONICHOOG.  
 DR ProDom: PD003042; HH\_signal.  
 DR SMART: SM00305; Hintc; 1.  
 DR SMART: SM00306; Hintn; 1.  
 DR SMART: SM00306; Hintn; 1.  
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KW Signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 23  
 FT CHAIN 24 409  
 FT CHAIN 24 197  
 FT SITE 197 198  
 FT SITE 197 198  
 FT SITE 267 267  
 FT ACT\_SITE 270 270  
 FT BINDING 197 197  
 FT DOMAIN 32 35  
 FT LIPID 24 24  
 SO SEQUENCE 409 AA; 4591 MW; 6FB265367FB98627 CRC64;

Query Match 55.48; Score 1267; DB 1; Length 409;  
 Best Local Similarity 59.9%; Pred. No. 2, 1e-94;  
 Matches 257; Conservative 49; Mismatches 95; Indels 28; Gaps 6;

Qy 5 LARCEVLIALSLVPCGL-ACGPGRGFRHRRKTLPLAYQGFIVNVEKTAGAGRY 63  
 Db 3 LPAVLLCAHALLSGAVKCGGRVYGRRRRTKLSPLSTYKQSPNVPEKTAGAGRY 62  
 Qy 64 EGIKTRNSERKELTPYNDIIFKDEENTGADRLTORCKDKLNALASVMNQPVRL 123  
 Db 63 EGIKTRNSERKELTPYNDIIFKDEITGADRLTORCKDKLNALASVMNQPVKL 122  
 Qy 124 RYEGWDEGHNSSESLHYEGRAVDITSDRSKYMLARLAVAGFDWYVESKAHII 183  
 Db 123 RYEGWDEGHNSSESLHYEGRAVDITSDRSKYMLARLAVAGFDWYVESKAHII 182  
 Qy 184 CSVAKESSVAAKSGCGPGSATVHLEGGKIVDLPGRVLAADOGRLYSDFLFL 243  
 Db 183 CSVAKESSVAAKSGCGPGSATVHLEGGKIVDLPGRVLAADOGRLYSDFLFL 242  
 Qy 244 DRDGAARKVYVETLEPRERLLTAHLFLVAPANDSCPTPGPSALFASVRPGQHYV 303  
 Db 243 DHPKKEHMOVITQOPRHRLFLTPAHLFLVS-DNYSTPASEQAVFASVVRGQYILV 301  
 Qy 304 VAEGGGRLLPAVHVTLEEEAGAVAPLTANGTILINVLASCVAVIEHSMHARAF 363  
 Db 302 SNVVG----LIPAVRVSVN-TQTYGAVAPLTQCTLVDDVVVSCALVQOKRLADIV 356  
 Qy 364 APFLAHLAALAPARTDGGGGSIPAQSAATEARGA-EPTAGIHWYSOLLVHIGTUL 423  
 Db 357 WPLRVLYNL-----GIAGTQPSQ-----MGIIHWYSKALYHIGRLIHL 395  
 Qy 424 SETMHPILGMA 432  
 Db 396 GNEFHPILGI 404

RESULT 13  
 IHH\_BRARE STANDARD; PRT; 412 AA.  
 ID IHH\_BRARE  
 AC Q98862;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Indian hedgehog protein precursor (IHH) (Fchidna hedgehog protein).

GN ITH OR EHH.  
 OS Brachydanio rerio (zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=96310864; PubMed=8684485;  
 RT Currie P.D., Ingham P.W.;  
 "Induction of a specific muscle cell type by a hedgehog-like protein  
 in zebrafish".  
 RL Nature 382:452-455(1996).  
 RN [2]  
 RP SEQUENCE OF 113-170 FROM N.A.  
 RC TISSUE=Muscle;  
 RX MEDLINE=97075114; PubMed=8917540;  
 RA Zardoya R., Abouheif E., Meyer A.;  
 "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species  
 closely related to the zebrafish".  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE  
 NOTOCHORD THAT INDUCES SOMITE PATTERNING AND MUSCLE PIONEER  
 DIFFERENTIATION.  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 TO MEDiate SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN THE NOTOCHORD.  
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT THE MID-GASTRULA STAGE.  
 CC -1- DISAPPEARS AT THE END OF THE SOMITOGENESIS.  
 CC -1- PPM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY  
 AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).  
 CC -----  
 DR EMBL: Y08426; CAA69702.1; -  
 DR HSSP: 062226; 1VHH.  
 DR MEROPS: C46 0PW: -  
 DR ZFIN: ZDB-GENE-980526-135; eh.  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hinc.  
 DR InterPro: IPR003586; Hedgehog\_hinc.  
 DR InterPro: IPR003587; Hedgehog\_hinc.  
 DR InterPro: IPR002203; Inteln.  
 DR InterPro: IPR001657; Son1cHH.  
 DR Pfam: PF01079; Hinc\_1.  
 DR Pfam: PF01085; HH\_signal\_1.  
 DR PRINTS: PR00632; SON1CHOG.  
 DR ProDom: PD003042; HH\_signal\_1.  
 DR SMART: SM00305; Hinc\_1.  
 DR SMART: SM00306; Hinc\_1.  
 DR PROSITE: PS50817; INTEIN\_N\_TER\_1.  
 KM Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 RN Signal: Lipoprotein; Palmitate.  
 FT SIGNAL 1 23  
 FT CHAIN 24 412  
 FT 197 INDIAN HEDGEHOG PROTEIN.  
 FT 24 INDIAN HEDGEHOG PROTEIN N-PRODUCT.

FT CHAIN 198 412 INDIAN HEDGEHOG PROTEIN C-PRODUCT.  
 FT SITE 197 198 CLEAVAGE (AUTO-) (BY SIMILARITY).  
 FT SITE 246 246 INVOLVED IN CHOLESTEROL TRANSFER (BY  
 FT SITE 269 269 INVOLVED IN AUTO-CLEAVAGE (BY  
 FT ACT\_SITE 272 272 SIMILARITY).  
 FT BINDING 197 197 ESSENTIAL FOR AUTO-CLEAVAGE (BY  
 FT LIPID 24 24 CHOLESTEROL (BY SIMILARITY).  
 FT CONFLICT 121 121 PALMITATE (BY SIMILARITY).  
 FT SEQUENCE 412 AA; 45746 MW; 21043F052A276CAL CRC64;  
 Query Match 53.6%; Score 1228; DB 1; Length 412;  
 Best Local Similarity 60.4%; Pred. No. 2,9e-91;  
 Matches 252; Conservative 43; Mismatches 80; Indels 42; Gaps 8;  
 QY 25 CGPGRGFGKRRHPKKTPLATYKOFITPNAVEKTLGASRGKTRNSRFPKLTPLYND 84  
 DB 24 CGPGRGFGKRRHPKKTPLATYKOFITPNAVEKTLGASRGKTRNSRFPKLTPLYND 83  
 QY 85 ITFKDEMTGADRLMTORCKDKLNALISVMNOMPGRVRLVTEGMDEGHSESLHYEG 144  
 DB 84 ITFKDEMTGADRLMTORCKDKLNALISVMNOMPGRVRLVTEGMDEGHSESLHYEG 143  
 QY 145 RAVDITTSRDRSKYGMRLAVEAGFDWYVESKAHICSVAKNSVAKSGGCPGSA 204  
 DB 144 RAVDITTSRDRSKYGMRLAVEAGFDWYVESKAHICSVAKNSVAKSGGCPGSA 203  
 QY 205 TYHLEGGTKLVKDLRFGRVLA--ADDOGRLLSYDFTFLDRDGAKKVFYIETLEP 261  
 DB 204 LYTKMGDSRQRLDQAGLVLAESGSDGTGLDYSEVLTFDRRPTQKHFEYVIRT-ED 262  
 QY 262 REBLITTAHLFLVAPNDSG--PRGP-SALFASRVPRGQVYVYVARGDRLPLPA 317  
 DB 263 GASVSLTAHLFLVAPNDSG--PRGP-SALFASRVPRGQVYVYVARGDRLPLPA 317  
 QY 318 VHSVTLREEGAYAPLTAHTGILLIRVLASCAVYEESMAHRAFPRLAHALALALA 377  
 DB 318 ITHVGRREQ-GLYPLPLTAHGVVVDVLTSCAAVNRRLAHMAFAPRLALYSW----- 371  
 QY 378 PARTDGGGGSTPAMASATEANGAEPFA--GIHWISQLIYHIGTWLDSETHPLGM 432  
 DB 372 -----TGPDQVLKNGLHWISQVILGKLLDSELPPLAL 407  
 RESULT 14  
 DH\_MOUSE STANDARD; PRT; 396 AA.  
 AC 061488;  
 DT 15-JUL-1999 (rel. 38, Created)  
 DT 15-JUL-1999 (rel. 38, Last sequence update)  
 DT 15-OCT-2001 (rel. 40, Last annotation update)  
 DE Desert hedgehog protein precursor (DHH) (HHG-3).  
 GN DHH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv;  
 RX MEDLINE=94094334; PubMed=7916661;  
 RA Echelard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,  
 McMahon J.A., McMahon A.P.;  
 "Sonic hedgehog, a member of a family of putative signaling  
 molecules, is implicated in the regulation of CNS polarity".  
 RT Cell 75:1417-1430(1993).  
 RL [2]  
 RP SEQUENCE OF 120-168 FROM N.A.  
 RX MEDLINE=95236997; PubMed=7720571;  
 RA Chang D.T., Lopez A., von Kessler A., Chiang C., Simandl B.K.,  
 Zhao R., Seldin M.F., Fallon J.F., Beachy P.A.;



RT "Products, genetic linkage and limb patterning activity of a murine  
 RT hedgehog gene.";  
 RL Development 120:3339-3353(1994).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A  
 CC SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES  
 CC DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTES. NOT EXPRESSED IN  
 CC LIMB BUDS.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 DR EMBL: X76292; CAA53924.1; -;  
 DR HSSP: O62226; LVH.  
 DR MEROPS: C46.004; -;  
 DR MKD: MGI:94891; Dh.  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hnt.  
 DR InterPro: IPR003586; Hedgehog\_hntc.  
 DR InterPro: IPR003587; Hedgehog\_hntn.  
 DR InterPro: IPR002203; Intein.  
 DR InterPro: IPR001657; SonichH.  
 DR Pfam: PF01079; Hntc; 1.  
 DR Pfam: PF01085; HH\_signal; 1.  
 DR PRINTS: PR00632; SONICHOG.  
 DR ProDom: PD003042; HH\_signal; 1.  
 DR SMART: SM00305; Hntc; 1.  
 DR SMART: SM00306; Hntn; 1.  
 DR PROSITE: PS50817; INTEIN\_N\_TER; 1.  
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KW Signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 22  
 FT CHAIN 23 396  
 FT CHAIN 23 198  
 FT CHAIN 199 396  
 FT SITE 198 199  
 FT SITE 244 244  
 FT SITE 244 244  
 FT SITE 268 268  
 FT ACT\_SITE 271 271  
 FT BINDING 198 198  
 FT LIPID 23 23  
 FT SEQUENCE 396 AA; 43542 MW; AFEE051B8950F8 CRC64;  
 Query Match 53.3%; Score 1219; DB 1; Length 396;  
 Best Local Similarity 58.7%; Pred. No. 1.5e-90;  
 Matches 249; Conservative 55; Mismatches 80; Indels 40; Gaps 12;

QY 62 RYEGKTRNSERFKEKLPNTNPDIIFDEENTGADRLMTORCKDKLNALISVNMOPGV 121  
 DB 62 PAECRVYRGSRFPDLPVNPNDIIFDEENSGADRLMTERCKRVNALIAVNMMPGV 121  
 QY 122 RLRTVETGDEDEGHSESLHYEGRAVDITTSDDRSTYGMALRLAVEAGDWMYTESKAH 161  
 DB 122 RLRTVETGDEDEGHSESLHYEGRAVDITTSDDRSTYGMALRLAVEAGDWMYTESKAH 161  
 QY 182 IHGCVKANSVAAKSGGCPGSAVHLEGGTKLVKLRGDRVLADDDGGRLLYSDFLT 241  
 DB 182 IHVSFKDNSLVARAGGCFEGNATVRLRSGERGLRSLHNGDVLADDAAGRVPTVLL 241  
 QY 242 FLRDEGAKYFYVYIETLEPRERILLTAHLLFYAPHNDSGPTPGP---SALFASRYRPG 298  
 DB 242 FLRDLDRASFPVAVETERRPRKLLPMLHVFNA---AGPAPQDFAPVFRARRLRAG 297  
 QY 299 QRYVVAERGGDRLLPAVHSTVLRREAGANAPLTAHGTILINRLVASCYAVIEHSHV 358  
 DB 298 DSVLA---PGGD-ALQPARVARVA-REAVGVFAPLTAHGTLLVNDVLASCYAVLESHOW 352  
 QY 359 AHRAPFRLAHLALAPARTGGGGSPAAQSTEARGAEPPTAGIHVYSLYHIG 418  
 DB 353 AHRAPFRLHLA-LGALLP-----GGAV-----QPT-GMHVYSLYRLA 391  
 QY 419 TWLL 422  
 DB 392 EELM 395  
 RESULT 15  
 ID DHH\_HUMAN STANDARD; PRT; 396 AA.  
 AC 043323; Q15794;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Desert hedgehog protein precursor (DHH) (HHG-3).  
 GN DHH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tate G., Kishimoto K., Mitsuura T.;  
 RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in  
 RT human cancer cell lines and embryonic organs.";  
 RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).  
 RN [2]  
 RP SEQUENCE OF 85-178 FROM N.A.  
 RA Tissue-Kidney;  
 RC Tissue-Kidney;  
 RA Drummond I.A.;  
 RT "Human desert hedgehog.";  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP DISEASE.  
 RX PubMed=11017805;  
 RA Umehara F., Tate G., Itoh K., Yamaguchi N., Douchi T., Mitsuura T.,  
 RA Osame M.;  
 RT "A novel mutation of desert hedgehog in a patient with 46,xy partial  
 RT gonadal dysgenesis accompanied by multifascicular neuropathy.";  
 RL Am. J. Hum. Genet. 67:1302-1305(2000).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A  
 CC SPERMATOCYTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES  
 CC DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN

CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESPECTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING. WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- DISEASE: Defects in DHH may be the cause of partial gonadal  
 CC dysgenesis (PGD) with minifascicular neuropathy. PGD is  
 CC characterized by the presence of a testis on one side and a streak  
 CC or an absent gonad at the other, persistence of Mullerian duct  
 CC structures, and a variable degree of genital ambiguity.  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AB010994; BAA24866.1; -  
 DR EMBL: AB010581; BAA24866.1; JOINED.  
 DR EMBL: AB010993; BAA24866.1; JOINED.  
 DR EMBL: 059748; AAB03398.1; -  
 DR HSP: 062226; 1VHH.  
 DR MEROPS: C46.004; -  
 DR GeneW: HGNC:2865; DHH.  
 DR MIM: 605423; -  
 DR MIM: 607080; -  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hinc.  
 DR InterPro: IPR003586; Hedgehog\_hinc.  
 DR InterPro: IPR003587; Hedgehog\_hinc.  
 DR InterPro: IPR001657; SonicHH.  
 DR Pfam: PF01079; Hhnt; 1.  
 DR Pfam: PF01085; HH\_signal; 1.  
 DR PRINTS: PRO0632; SONICHHOG.  
 DR ProDom: PD003042; HH\_signal; 1.  
 DR SMART: SM00305; Hhnt; 1.  
 DR SMART: SM00306; Hhnt; 1.  
 DR SMART: SM00306; Hhnt; 1.  
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 DR Signal; Lipoprotein; Palmitate.  
 KM Signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 22  
 FT CHAIN 23 396  
 FT CHAIN 23 198  
 FT CHAIN 199 396  
 FT CHAIN 198 199  
 FT SITE 244 244  
 FT SITE 244 244  
 FT SITE 268 268  
 FT SITE 268 268  
 FT ACT\_SITE 271 271  
 FT BINDING 198 198  
 FT LIPID 23 23  
 FT CONFLICT 177 177  
 FT CONFLICT 177 177  
 SQ SEQUENCE 396 AA; 43577 MW; FCE4FB21972C3AD5 CRC64;

Query Match 53.2%; Score 1217; DB 1; Length 396;  
 Best Local Similarity 58.7%; Pred. No. 2,1e-90;  
 Matches 249; Conservative 53; Mismatches 82; Indels 40; Gaps 12;

4 LARCFVLIVASSLVCPLACGPRG-GKKRRHPKK-LTPLAYKQIFPNAVEKTLGASG 61  
 7 LPLPLCLALLA-----LPQSCGPGRGPGVGRRYARKQLVPLLYKQFVGPVPERTIGASG 61  
 62 RYEGKITRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLALAIQVNMQPGV 121  
 62 PABGRVARGSERPRDLVPYNPDIIFKDEENGADRLMTORCKDKLALAIQVNMQPGV 121  
 122 RLRTVEGMDEGDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYESKAH 181

DB 122 RLRTVEGMDEGDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVEAGFDWVYESKAH 181  
 182 IHOSVKAENSVAKSGCGCPGSAATVLEQCGTFLVDLRPGDRLAADDGRLYSDFLT 241  
 182 VHSYVAADNSLAVRAGCGCPGNATVTLMSGERKGLNELRGDVLADASGRVPTPVLL 241  
 242 FLDRDGAARKVYVITLPPRERLITTAHLLEFVAPHNSGPTPGP---SALFASVRPG 298  
 242 FLDRDQRRASFAVETEMPRLKLLTPMHLVFAA----RGPARPGDFAPVAPRLRAG 297  
 299 QRYTVVAENGDRRLPAVHSYTLREERAGAVPLTANGTILINRYLASCAVIEESHW 358  
 298 DSVLA---PGGD-ALRPARVARVA-REEAVGVFAPLTAHGTILVNDYLASCAVIESHOW 352  
 359 AHRAPFRLAHLAALAPARTDGGGSIIPAQASATARGAEPYAGIHWYSOLLYHTG 418  
 353 AHRAPFRLAHLA-LEALLP-----GGAV-----QPT-GMHWYSLYRLA 391  
 419 TWLL 422  
 392 EELL 395

Search completed: February 20, 2003, 10:12:39  
 Job time : 15.409 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:10:28 ; Search time 7.84443 Seconds  
(without alignments)  
1423.285 Million cell updates/sec

Title: US-09-827-110A-13

Perfect score: 2289 1 MLILARCLVLILASSLVC.....GTWLIDSETHPIGLMAVKSS 437

Sequence:

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 2554876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB pep:\*

2: /cgn2\_6/ptodata/1/pubppaa/PCr\_NEW\_PUB pep:\*

3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB pep:\*

4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB pep:\*

5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB pep:\*

6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB pep:\*

7: /cgn2\_6/ptodata/1/pubppaa/PCrUS\_PUBCOMB pep:\*

8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB pep:\*

9: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB pep:\*

10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB pep:\*

11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB pep:\*

12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB pep:\*

13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB pep:\*

14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2289	100.0	437	8 US-08-900-220C-13	Sequence 13, Appl
2	2289	100.0	437	9 US-09-883-848A-13	Sequence 13, Appl
3	2289	100.0	437	10 US-09-151-999-13	Sequence 13, Appl
4	2286	99.9	437	9 US-09-950-046-14	Sequence 14, Appl
5	2286	99.9	437	9 US-09-969-520A-1	Sequence 1, Appl
6	2286	99.9	437	9 US-10-013-310-1	Sequence 16, Appl
7	2286	99.9	437	9 US-09-733-634-16	Sequence 37, Appl
8	2286	99.9	437	10 US-09-021-660A-37	Sequence 3, Appl
9	2280	99.6	437	9 US-09-969-520A-3	Sequence 10, Appl
10	2277	99.5	437	9 US-09-969-520A-10	Sequence 11, Appl
11	2277	99.5	437	9 US-09-969-520A-11	Sequence 2, Appl
12	2276	99.4	437	9 US-08-968-520A-2	Sequence 4, Appl
13	2272	99.3	437	9 US-09-969-520A-9	Sequence 8, Appl
14	2270	99.2	437	9 US-09-969-520A-4	Sequence 7, Appl
15	2266	99.0	437	9 US-09-969-520A-6	Sequence 5, Appl
16	2263	98.9	437	9 US-09-969-520A-7	Sequence 14, Appl
17	2254	98.5	437	9 US-09-969-520A-5	Sequence 14, Appl
18	2237	97.7	437	9 US-09-733-634-14	Sequence 14, Appl
19	2065	90.2	462	9 US-09-883-848A-15	Sequence 15, Appl

20	2065	90.2	475	8 US-08-900-220C-15	Sequence 15, Appl
21	2065	90.2	475	9 US-09-883-848A-15	Sequence 15, Appl
22	2065	90.2	475	10 US-09-021-660A-39	Sequence 39, Appl
23	2065	90.2	475	10 US-09-151-999-15	Sequence 15, Appl
24	1788	78.1	425	8 US-08-900-220C-10	Sequence 10, Appl
25	1788	78.1	425	8 US-08-883-848A-10	Sequence 10, Appl
26	1788	78.1	425	10 US-09-021-660A-34	Sequence 34, Appl
27	1788	78.1	425	10 US-09-151-999-10	Sequence 10, Appl
28	1517	66.3	418	10 US-09-021-660A-38	Sequence 38, Appl
29	1466	64.0	418	8 US-08-900-220C-14	Sequence 14, Appl
30	1466	64.0	418	9 US-08-883-848A-14	Sequence 14, Appl
31	1466	64.0	418	10 US-09-151-999-14	Sequence 14, Appl
32	1464	64.0	416	8 US-08-900-220C-18	Sequence 18, Appl
33	1464	64.0	416	8 US-08-883-848A-18	Sequence 18, Appl
34	1464	64.0	416	10 US-09-151-999-18	Sequence 18, Appl
35	1276.5	55.8	411	8 US-08-900-220C-16	Sequence 16, Appl
36	1276.5	55.8	411	9 US-08-883-848A-16	Sequence 16, Appl
37	1276.5	55.8	411	10 US-09-151-999-16	Sequence 16, Appl
38	1271.5	55.5	411	9 US-09-733-634-28	Sequence 28, Appl
39	1268	55.4	449	9 US-09-990-046-29	Sequence 29, Appl
40	1268	55.4	449	9 US-08-900-220C-12	Sequence 30, Appl
41	1240	54.2	411	8 US-08-900-220C-12	Sequence 12, Appl
42	1240	54.2	411	9 US-08-883-848A-12	Sequence 12, Appl
43	1240	54.2	411	10 US-09-151-999-12	Sequence 12, Appl
44	1219	53.3	396	8 US-08-900-220C-11	Sequence 11, Appl
45	1219	53.3	396	9 US-09-990-046-13	Sequence 13, Appl

#### ALIGNMENTS

RESULT 1  
US-08-900-220C-13  
Sequence 13, Application US/08900220C  
Patent No. US20020045206A1

GENERAL INFORMATION:  
APPLICANT: Miao, Ningning  
Wang, Monica  
Mahantappa, Nagesh K.  
Pang, Kevin  
Jin, Ping

TITLE OF INVENTION: Method of Treating Dopaminergic and  
GABA-ergic Disorders

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELLIOT LLP  
STREET: ONE POST OFFICE SQUARE  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Ascii (text)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/900,220C  
FILING DATE: 24-JUL-1997  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: ONV-044.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-7000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-08-900-220C-13

## Query Match

Best Local Similarity 100.0%; Score 2289; DB 8; Length 437;  
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPPLAYKQFIPNVAEKTIGAS 60
QY 61 GRYEGKITRNSERFKELTLPNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
DB 61 GRYEGKITRNSERFKELTLPNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
QY 121 VRLRTGEMDEDDGHSESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYYESKA 180
DB 121 VRLRTGEMDEDDGHSESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYYESKA 180
QY 181 HHSVKAENSVAAKSGCGCPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
DB 181 HHSVKAENSVAAKSGCGCPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
QY 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDSGPTPGPSALFASRVPRGQR 300
DB 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDSGPTPGPSALFASRVPRGQR 300
QY 301 VYVVAERGDRRLPLAAVHSTVTLREBEAGAVAPLTAGTILINRVLASCYAVIEESHMAH 360
DB 301 VYVVAERGDRRLPLAAVHSTVTLREBEAGAVAPLTAGTILINRVLASCYAVIEESHMAH 360
QY 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPITAGIHMYSQLYHIGTW 420
DB 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPITAGIHMYSQLYHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

## RESULT 2

US-09-883-848A-13

```

; Sequence 13, Application US/09883848A
; Publication No. US20030022819A1
; GENERAL INFORMATION:
; APPLICANT: Ling, L.
; TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES
; FILE REFERENCE: CIBT-P01-119
; CURRENT APPLICATION NUMBER: US/09/883,848A
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/211,919
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-883-848A-13

```

Query Match 100.0%; Score 2289; DB 9; Length 437;  
Best Local Similarity 100.0%; Pred. No. 1.2e-197;  
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPPLAYKQFIPNVAEKTIGAS 60
QY 61 GRYEGKITRNSERFKELTLPNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
DB 61 GRYEGKITRNSERFKELTLPNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120

```

```

QY 121 VRLRTGEMDEDDGHSESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYYESKA 180
DB 121 VRLRTGEMDEDDGHSESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYYESKA 180
QY 181 HHSVKAENSVAAKSGCGCPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
DB 181 HHSVKAENSVAAKSGCGCPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
QY 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDSGPTPGPSALFASRVPRGQR 300
DB 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDSGPTPGPSALFASRVPRGQR 300
QY 301 VYVVAERGDRRLPLAAVHSTVTLREBEAGAVAPLTAGTILINRVLASCYAVIEESHMAH 360
DB 301 VYVVAERGDRRLPLAAVHSTVTLREBEAGAVAPLTAGTILINRVLASCYAVIEESHMAH 360
QY 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPITAGIHMYSQLYHIGTW 420
DB 361 RAFAFRLAHALLAALAPARTDGGGGSIPAAQSATEARGAEPITAGIHMYSQLYHIGTW 420
QY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

## RESULT 3

US-09-151-999-13

```

; Sequence 13, Application US/09151999
; Patent No. US20020151460A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Elizabeth
; TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE
; FILE REFERENCE: ONV-031.02
; CURRENT APPLICATION NUMBER: US/09/151,999
; EARLIER FILING DATE: 1998-08-11
; EARLIER APPLICATION NUMBER: 08/955,552
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 437
; TYPE: PRT
; ORGANISM: murine Shh
US-09-151-999-13

```

Query Match 100.0%; Score 2289; DB 10; Length 437;  
Best Local Similarity 100.0%; Pred. No. 1.2e-197;  
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFRGRRHPKKTLPPLAYKQFIPNVAEKTIGAS 60
QY 61 GRYEGKITRNSERFKELTLPNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
DB 61 GRYEGKITRNSERFKELTLPNNPDIIFKDEENTGADRLMTORCKDKLNALAI SVNMWPG 120
QY 121 VRLRTGEMDEDDGHSESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYYESKA 180
DB 121 VRLRTGEMDEDDGHSESLHYEGRAVDITTSDRDSKYGMLARLAVEAGFDWYYESKA 180
QY 181 HHSVKAENSVAAKSGCGCPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
DB 181 HHSVKAENSVAAKSGCGCPGSATVHLEGGGTRKLVKDLRPGDRVLAADDOGRLLYSDFL 240
QY 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDSGPTPGPSALFASRVPRGQR 300
DB 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLTFVAPHNDSGPTPGPSALFASRVPRGQR 300
QY 301 VYVVAERGDRRLPLAAVHSTVTLREBEAGAVAPLTAGTILINRVLASCYAVIEESHMAH 360
DB 301 VYVVAERGDRRLPLAAVHSTVTLREBEAGAVAPLTAGTILINRVLASCYAVIEESHMAH 360

```

QY 361 RAFAFRLAALLAALAPARTDGGGSIIPAASATAEARGAEPTAGIHWXSOLLHIGTW 420  
DB 361 RAFAFRLAALLAALAPARTDGGGSIIPAASATAEARGAEPTAGIHWXSOLLHIGTW 420  
QY 421 LLDSETHMPLGMAVKSS 437  
DB 421 LLDSETHMPLGMAVKSS 437

RESULT 4  
US-09-990-046-14  
; Sequence 14, Application US/09990046  
; Patent No. US20020156245A1  
; GENERAL INFORMATION:  
; APPLICANT: de Sauvage, Frederic  
; APPLICANT: Carpenter, David A.  
; TITLE OF INVENTION: Patched-2  
; FILE REFERENCE: P1405R1  
; CURRENT APPLICATION NUMBER: US/09/990,046  
; PRIOR FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/293,505  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-15  
; NUMBER OF SEQ ID NOS: 32  
; SEQ ID NO 14  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-990-046-14

Query Match 99.9%; Score 2286; DB 9; Length 437;  
Best Local Similarity 99.8%; Pred. No. 2.3e-197;  
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLLARCFVLIASSILVCPGLACGPGRGFGRRHKITPLAYKQFTINVAEKTIGAS 60  
DB 1 MLLLLARCFVLIASSILVCPGLACGPGRGFGRRHKITPLAYKQFTINVAEKTIGAS 60  
QY 61 GREGKITRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNOMP 120  
DB 61 GREGKITRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNOMP 120  
QY 121 VRLVTEGMEDEGHSESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDWVYESKA 180  
DB 121 VRLVTEGMEDEGHSESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDWVYESKA 180  
QY 181 HHCISVKAENSVAAGSGCCPGSATVHLEGGTKLVKDLRPGGRVLAADOGRLYSDFL 240  
DB 181 HHCISVKAENSVAAGSGCCPGSATVHLEGGTKLVKDLRPGGRVLAADOGRLYSDFL 240  
QY 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAAPHNDGPTPGPSALFASRVPRGQR 300  
DB 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAAPHNDGPTPGPSALFASRVPRGQR 300  
QY 301 VYVVAERGGDRLLPAAVHSVTLREEBAGAYADPLTAHGTLLINRVLASCAVIEESHMAH 360  
DB 301 VYVVAERGGDRLLPAAVHSVTLREEBAGAYADPLTAHGTLLINRVLASCAVIEESHMAH 360  
QY 361 RAFAFRLAALLAALAPARTDGGGSIIPAASATAEARGAEPTAGIHWXSOLLHIGTW 420  
DB 361 RAFAFRLAALLAALAPARTDGGGSIIPAASATAEARGAEPTAGIHWXSOLLHIGTW 420  
QY 421 LLDSETHMPLGMAVKSS 437  
DB 421 LLDSETHMPLGMAVKSS 437

RESULT 5  
US-09-969-520A-1  
; Sequence 1, Application US/09969520A  
; Patent No. US20020177163A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE

; APPLICANT: BEACHY, Philip A.  
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHE  
; FILE REFERENCE: JHU1670-1  
; CURRENT APPLICATION NUMBER: US/09/969,520A  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-969-520A-1

Query Match 99.9%; Score 2286; DB 9; Length 437;  
Best Local Similarity 99.8%; Pred. No. 2.3e-197;  
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLLLARCFVLIASSILVCPGLACGPGRGFGRRHKITPLAYKQFTINVAEKTIGAS 60  
DB 1 MLLLLARCFVLIASSILVCPGLACGPGRGFGRRHKITPLAYKQFTINVAEKTIGAS 60  
QY 61 GREGKITRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNOMP 120  
DB 61 GREGKITRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALAI SVMNOMP 120  
QY 121 VRLVTEGMEDEGHSESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDWVYESKA 180  
DB 121 VRLVTEGMEDEGHSESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDWVYESKA 180  
QY 181 HHCISVKAENSVAAGSGCCPGSATVHLEGGTKLVKDLRPGGRVLAADOGRLYSDFL 240  
DB 181 HHCISVKAENSVAAGSGCCPGSATVHLEGGTKLVKDLRPGGRVLAADOGRLYSDFL 240  
QY 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAAPHNDGPTPGPSALFASRVPRGQR 300  
DB 241 TFLDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAAPHNDGPTPGPSALFASRVPRGQR 300  
QY 301 VYVVAERGGDRLLPAAVHSVTLREEBAGAYADPLTAHGTLLINRVLASCAVIEESHMAH 360  
DB 301 VYVVAERGGDRLLPAAVHSVTLREEBAGAYADPLTAHGTLLINRVLASCAVIEESHMAH 360  
QY 361 RAFAFRLAALLAALAPARTDGGGSIIPAASATAEARGAEPTAGIHWXSOLLHIGTW 420  
DB 361 RAFAFRLAALLAALAPARTDGGGSIIPAASATAEARGAEPTAGIHWXSOLLHIGTW 420  
QY 421 LLDSETHMPLGMAVKSS 437  
DB 421 LLDSETHMPLGMAVKSS 437

RESULT 6  
US-10-013-310-1  
; Sequence 1, Application US/10013310  
; Publication No. US20020192216A1  
; GENERAL INFORMATION:  
; APPLICANT: Lamb, Jonathan Robert  
; APPLICANT: Hoynes, Gerard Francis  
; APPLICANT: Dallman, Margaret Jane  
; TITLE OF INVENTION: Therapeutic Use  
; FILE REFERENCE: 674525-2003  
; CURRENT APPLICATION NUMBER: US/10/013,310  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR FILING DATE: PCT/GB00/02191  
; PRIOR FILING DATE: 2000-06-05  
; PRIOR APPLICATION NUMBER: UK 9913350.6  
; PRIOR FILING DATE: 1999-06-08  
; PRIOR APPLICATION NUMBER: UK 9921953.7  
; PRIOR FILING DATE: 1999-09-16  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1

```

; LENGTH: 437
; TYPE: PRT
; ORGANISM: House Mouse
US-10-013-310-1

```

```

Query Match          99.9%: Score 2286; DB 9; Length 437;
Beet Local Similarity 99.8%: Pred. No. 2.3e-197;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
OY 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTORCKKLNLAISVNMOWPG 120
DB 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTORCKKLNLAISVNMOWPG 120
OY 121 VRLRTEGMDGDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
DB 121 VRLRTEGMDGDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
OY 181 HHCYKAENSVAAKSGCGPFSATVHLEOGGTYKLVKDLRPGDRVLAADDGRLLYSDFL 240
DB 181 HHCYKAENSVAAKSGCGPFSATVHLEOGGTYKLVKDLRPGDRVLAADDGRLLYSDFL 240
OY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
OY 301 VYVAERGGDRLLPAAVHSTVTLREBEAGAYAPLTAHGTTILINRVLASCYAVIEESHMAH 360
DB 301 VYVAERGGDRLLPAAVHSTVTLREBEAGAYAPLTAHGTTILINRVLASCYAVIEESHMAH 360
OY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPPIAGIMYSOLLHYIGTW 420
DB 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPPIAGIMYSOLLHYIGTW 420
OY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

```

RESULT 7
US-09-733-634-16
; Sequence 16, Application US/09733634
; Publication No. US20030013646A1
; GENERAL INFORMATION:
; APPLICANT: Massachusetts General Hospital
; TITLE OF INVENTION: Method to stimulate Insulin production by pancreatic b-cells
; FILE REFERENCE: 17633/1240
; CURRENT APPLICATION NUMBER: US/09/733,634
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/170,282
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentn version 3.0
; SEQ ID NO 16
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-733-634-16

```

```

Query Match          99.9%: Score 2286; DB 9; Length 437;
Beet Local Similarity 99.8%: Pred. No. 2.3e-197;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
OY 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTORCKKLNLAISVNMOWPG 120
DB 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTORCKKLNLAISVNMOWPG 120
OY 121 VRLRTEGMDGDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
DB 121 VRLRTEGMDGDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
OY 181 HHCYKAENSVAAKSGCGPFSATVHLEOGGTYKLVKDLRPGDRVLAADDGRLLYSDFL 240
DB 181 HHCYKAENSVAAKSGCGPFSATVHLEOGGTYKLVKDLRPGDRVLAADDGRLLYSDFL 240
OY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300

```

```

OY 121 VRLRTEGMDGDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
DB 121 VRLRTEGMDGDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
OY 181 HHCYKAENSVAAKSGCGPFSATVHLEOGGTYKLVKDLRPGDRVLAADDGRLLYSDFL 240
DB 181 HHCYKAENSVAAKSGCGPFSATVHLEOGGTYKLVKDLRPGDRVLAADDGRLLYSDFL 240
OY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
OY 301 VYVAERGGDRLLPAAVHSTVTLREBEAGAYAPLTAHGTTILINRVLASCYAVIEESHMAH 360
DB 301 VYVAERGGDRLLPAAVHSTVTLREBEAGAYAPLTAHGTTILINRVLASCYAVIEESHMAH 360
OY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPPIAGIMYSOLLHYIGTW 420
DB 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPPIAGIMYSOLLHYIGTW 420
OY 421 LLDSETHMPLGMAVKSS 437
DB 421 LLDSETHMPLGMAVKSS 437

```

```

RESULT 8
US-09-021-660A-37
; Sequence 37, Application US/09021660A
; Patent No. US20010041668A1
; GENERAL INFORMATION:
; APPLICANT: Baron, M.
; APPLICANT: Balausoff, M.
; TITLE OF INVENTION: METHODS FOR MODULATING HEMATOPOIESIS AND VASCULAR
; FILE REFERENCE: GROWTH
; CURRENT APPLICATION NUMBER: US/09/021,660A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/037,513
; PRIOR FILING DATE: 1997-02-10
; PRIOR APPLICATION NUMBER: 60/049,763
; PRIOR FILING DATE: 1997-06-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-021-660A-37

```

```

Query Match          99.9%: Score 2286; DB 10; Length 437;
Beet Local Similarity 99.8%: Pred. No. 2.3e-197;
Matches 436; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
DB 1 MLLLLARCFVLIVLASSLLVCPGLACGPGRGFGKRRHPKLTPLAYKQFIPNVAEKTIGAS 60
OY 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTORCKKLNLAISVNMOWPG 120
DB 61 GYEGKITRNSERFELPNYNPDIIFKDEENTGADRLMTORCKKLNLAISVNMOWPG 120
OY 121 VRLRTEGMDGDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
DB 121 VRLRTEGMDGDGHSESLHYEGRAVDITTSDRDSKYGMRLARLAVAGFDMVYTESKA 180
OY 181 HHCYKAENSVAAKSGCGPFSATVHLEOGGTYKLVKDLRPGDRVLAADDGRLLYSDFL 240
DB 181 HHCYKAENSVAAKSGCGPFSATVHLEOGGTYKLVKDLRPGDRVLAADDGRLLYSDFL 240
OY 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300
DB 241 TELDRDEGAKKVFYIETLEPRERLLTAAHLLEFVAPHNDGPTPGPSALFASRYRPGQR 300

```

Db 241 TFLDRBGAKKVFYVETLEPRERLLTAAHLFVAPHNDSGTPGPSALFASRVPRGQR 300  
Oy 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYAEUTAGHTLLINRVLASCAVIEESHMAH 360  
Db 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYAEUTAGHTLLINRVLASCAVIEESHMAH 360  
Oy 361 RAFAPRRLAALLAALAPARTDGGGGGSIIPAASATEARGAEPTAGIHWYSOLLYHIGTW 420  
Db 361 RAFAPRRLAALLAALAPARTDGGGGGSIIPAASATEARGAEPTAGIHWYSOLLYHIGTW 420  
Oy 421 LLDSETHPLGMAVKKSS 437  
Db 421 LLDSETHPLGMAVKKSS 437

RESULT 9  
US-09-969-520A-3  
; Sequence 3, Application US/09969520A  
; Patent No. US20020177163A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: BEACHY, Phillip A.  
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED  
; FILE REFERENCE: JHU1670-1  
; CURRENT APPLICATION NUMBER: US/09/969,520A  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 60/235,153  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 437  
; TYPE: PRP  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Altered sonic hedgehog protein  
US-09-969-520A-3

Query Match 99.6%; Score 2280; DB 9; Length 437;  
Best local Similarity 99.5%; Pred. No. 7.9e-197;  
Matches 435; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MLLLARCFVLIASSLVCPLGACGPGRGFGRRHPKKTPLATYKOFIPNVAEKTIGAS 60  
Db 1 MLLLARCFVLIASSLVCPLGACGPGRGFGRRHPKKTPLATYKOFIPNVAEKTIGAS 60  
Oy 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNOMP 120  
Db 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNOMP 120  
Oy 121 VRLVTEGMDGDHSESLHYEGRAVDITTSRDSRSKYGMLARLAVEAGFDWVYVESKA 180  
Db 121 VRLVTEGMDGDHSESLHYEGRAVDITTSRDSRSKYGMLARLAVEAGFDWVYVESKA 180  
Oy 181 HHCYSKAENSVAAKSGGCGPGSATYHLEGGGTKLYKDLRPGDRVLAADOGRLYSDFL 240  
Db 181 HHCYSKAENSVAAKSGGCGPGSATYHLEGGGTKLYKDLRPGDRVLAADOGRLYSDFL 240  
Oy 241 TFLDRBGAKKVFYVETLEPRERLLTAAHLFVAPHNDSGTPGPSALFASRVPRGQR 300  
Db 241 TFLDRBGAKKVFYVETLEPRERLLTAAHLFVAPHNDSGTPGPSALFASRVPRGQR 300  
Oy 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYAEUTAGHTLLINRVLASCAVIEESHMAH 360  
Db 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYAEUTAGHTLLINRVLASCAVIEESHMAH 360  
Oy 361 RAFAPRRLAALLAALAPARTDGGGGGSIIPAASATEARGAEPTAGIHWYSOLLYHIGTW 420  
Db 361 RAFAPRRLAALLAALAPARTDGGGGGSIIPAASATEARGAEPTAGIHWYSOLLYHIGTW 420  
Oy 421 LLDSETHPLGMAVKKSS 437  
Db 421 LLDSETHPLGMAVKKSS 437

RESULT 10  
US-09-969-520A-10  
; Sequence 10, Application US/09969520A  
; Patent No. US20020177163A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: BEACHY, Phillip A.  
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHE  
; FILE REFERENCE: JHU1670-1  
; CURRENT APPLICATION NUMBER: US/09/969,520A  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 60/235,153  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 437  
; TYPE: PRP  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Altered sonic hedgehog protein  
US-09-969-520A-10

Query Match 99.5%; Score 2277; DB 9; Length 437;  
Best local Similarity 99.3%; Pred. No. 1.5e-196;  
Matches 434; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MLLLARCFVLIASSLVCPLGACGPGRGFGRRHPKKTPLATYKOFIPNVAEKTIGAS 60  
Db 1 MLLLARCFVLIASSLVCPLGACGPGRGFGRRHPKKTPLATYKOFIPNVAEKTIGAS 60  
Oy 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNOMP 120  
Db 61 GREGKTRNSERFKELTPYNPDIIFKDEENTGADRLMTORCKDKLNALATISVMNOMP 120  
Oy 121 VRLVTEGMDGDHSESLHYEGRAVDITTSRDSRSKYGMLARLAVEAGFDWVYVESKA 180  
Db 121 VRLVTEGMDGDHSESLHYEGRAVDITTSRDSRSKYGMLARLAVEAGFDWVYVESKA 180  
Oy 181 HHCYSKAENSVAAKSGGCGPGSATYHLEGGGTKLYKDLRPGDRVLAADOGRLYSDFL 240  
Db 181 HHCYSKAENSVAAKSGGCGPGSATYHLEGGGTKLYKDLRPGDRVLAADOGRLYSDFL 240  
Oy 241 TFLDRBGAKKVFYVETLEPRERLLTAAHLFVAPHNDSGTPGPSALFASRVPRGQR 300  
Db 241 TFLDRBGAKKVFYVETLEPRERLLTAAHLFVAPHNDSGTPGPSALFASRVPRGQR 300  
Oy 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYAEUTAGHTLLINRVLASCAVIEESHMAH 360  
Db 301 VYVAERGGDRRLPAVAHSVTLREBEAGAYAEUTAGHTLLINRVLASCAVIEESHMAH 360  
Oy 361 RAFAPRRLAALLAALAPARTDGGGGGSIIPAASATEARGAEPTAGIHWYSOLLYHIGTW 420  
Db 361 RAFAPRRLAALLAALAPARTDGGGGGSIIPAASATEARGAEPTAGIHWYSOLLYHIGTW 420  
Oy 421 LLDSETHPLGMAVKKSS 437  
Db 421 LLDSETHPLGMAVKKSS 437

RESULT 11  
US-09-969-520A-11  
; Sequence 11, Application US/09969520A  
; Patent No. US20020177163A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: BEACHY, Phillip A.  
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHE  
; FILE REFERENCE: JHU1670-1  
; CURRENT APPLICATION NUMBER: US/09/969,520A  
; CURRENT FILING DATE: 2002-06-04

;; PRIOR APPLICATION NUMBER: US 60/235,153  
;; PRIOR FILING DATE: 2000-09-22  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 11  
;; LENGTH: 437  
;; TYPE: PRF  
;; ORGANISM: Artificial sequence  
;; FEATURE:  
;; OTHER INFORMATION: Altered sonic hedgehog protein  
US-09-969-520A-11

Query Match 99.5%; Score 2272; DB 9; Length 437;  
Best Local Similarity 99.3%; Pred. No. 1.5e-196;  
Matches 434; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLLVCGIACGPGRGFGKRRHPKRLTPLAYKQFIPVNAEKTIGAS 60  
DB 1 MLLIARCFVLIASSLLVCGIACGPGRGFGKRRHPKRLTPLAYKQFIPVNAEKTIGAS 60  
QY 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SYMNQMPG 120  
DB 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SYMNQMPG 120  
QY 121 VRLRVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGDWMYYESKA 180  
DB 121 VRLRVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGDWMYYESKA 180  
QY 181 HHCYKAENSVAKSGCGPGSATVHLEOGGKTLVKDLRPGDRLVLAADOGRLYSDFL 240  
DB 181 HHCYKAENSVAKSGCGPGSATVHLEOGGKTLVKDLRPGDRLVLAADOGRLYSDFL 240  
QY 241 TELDRDEGAKKVFYETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300  
DB 241 TELDRDEGAKKVFYETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300  
QY 301 VYVAERGGDRRLPAVHSVTLREBAGAYAPLTAHGTLLINRVLASCYAVEEHSMAH 360  
DB 301 VYVAERGGDRRLPAVHSVTLREBAGAYAPLTAHGTLLINRVLASCYAVEEHSMAH 360  
QY 361 RAFAFRLHALLAALAPARTDGGGGSIPAAQSATEARGAEPTAGIHWYSQLLYHGTW 420  
DB 361 RAFAFRLHALLAALAPARTDGGGGSIPAAQSATEARGAEPTAGIHWYSQLLYHGTW 420  
QY 421 LLDSETHMPLGMAVKSS 437  
DB 421 LLDSETHMPLGMAVKSS 437

## RESULT 12

US-09-969-520A-2  
;; Sequence 2, Application US/09969520A  
;; Patent No. US20020177163A1  
;; GENERAL INFORMATION:  
;; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
;; APPLICANT: BEACHY, Philip A.  
;; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED  
;; FILE REFERENCE: JH01670-1  
;; CURRENT APPLICATION NUMBER: US/09/969, 520A  
;; PRIOR FILING DATE: 2002-06-04  
;; PRIOR APPLICATION NUMBER: US 60/235,153  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 2  
;; LENGTH: 437  
;; TYPE: PRF  
;; ORGANISM: Artificial sequence  
;; FEATURE:  
;; OTHER INFORMATION: Altered sonic hedgehog protein  
US-09-969-520A-2

Query Match 99.4%; Score 2276; DB 9; Length 437;

Best Local Similarity 99.5%; Pred. No. 1.8e-196;  
Matches 435; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLLVCGIACGPGRGFGKRRHPKRLTPLAYKQFIPVNAEKTIGAS 60  
DB 1 MLLIARCFVLIASSLLVCGIACGPGRGFGKRRHPKRLTPLAYKQFIPVNAEKTIGAS 60  
QY 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SYMNQMPG 120  
DB 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SYMNQMPG 120  
QY 121 VRLRVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGDWMYYESKA 180  
DB 121 VRLRVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGDWMYYESKA 180  
QY 181 HHCYKAENSVAKSGCGPGSATVHLEOGGKTLVKDLRPGDRLVLAADOGRLYSDFL 240  
DB 181 HHCYKAENSVAKSGCGPGSATVHLEOGGKTLVKDLRPGDRLVLAADOGRLYSDFL 240  
QY 241 TELDRDEGAKKVFYETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300  
DB 241 TELDRDEGAKKVFYETLEPRERLLTAHLLFVAPHNDSGPTPGPSALFASRVPRGQR 300  
QY 301 VYVAERGGDRRLPAVHSVTLREBAGAYAPLTAHGTLLINRVLASCYAVEEHSMAH 360  
DB 301 VYVAERGGDRRLPAVHSVTLREBAGAYAPLTAHGTLLINRVLASCYAVEEHSMAH 360  
QY 361 RAFAFRLHALLAALAPARTDGGGGSIPAAQSATEARGAEPTAGIHWYSQLLYHGTW 420  
DB 361 RAFAFRLHALLAALAPARTDGGGGSIPAAQSATEARGAEPTAGIHWYSQLLYHGTW 420  
QY 421 LLDSETHMPLGMAVKSS 437  
DB 421 LLDSETHMPLGMAVKSS 437

## RESULT 13

US-09-969-520A-9  
;; Sequence 9, Application US/09969520A  
;; Patent No. US20020177163A1  
;; GENERAL INFORMATION:  
;; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
;; APPLICANT: BEACHY, Philip A.  
;; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED  
;; FILE REFERENCE: JH01670-1  
;; CURRENT APPLICATION NUMBER: US/09/969, 520A  
;; PRIOR FILING DATE: 2002-06-04  
;; PRIOR APPLICATION NUMBER: US 60/235,153  
;; NUMBER OF SEQ ID NOS: 17  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 9  
;; LENGTH: 437  
;; TYPE: PRF  
;; ORGANISM: Artificial sequence  
;; FEATURE:  
;; OTHER INFORMATION: Altered sonic hedgehog protein  
US-09-969-520A-9

Query Match 99.3%; Score 2272; DB 9; Length 437;  
Best Local Similarity 99.3%; Pred. No. 4.1e-196;  
Matches 434; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLLIARCFVLIASSLLVCGIACGPGRGFGKRRHPKRLTPLAYKQFIPVNAEKTIGAS 60  
DB 1 MLLIARCFVLIASSLLVCGIACGPGRGFGKRRHPKRLTPLAYKQFIPVNAEKTIGAS 60  
QY 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SYMNQMPG 120  
DB 61 GREGKITRNSERKELTPNPNPDIIKDEENTGADRLMTORCKDKLNALAI SYMNQMPG 120  
QY 121 VRLRVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGDWMYYESKA 180  
DB 121 VRLRVTEGMDGHHSESLHYEGRAVDITTSDDRSKYGMLARLAVAGDWMYYESKA 180

Db 121 VKLAVTEGMDDEGHHSESLHYEGRAVDITTSDDRKYKGLAFLAAYEAGFDWYYYESKA 180  
QY 181 HHCSSVAENSVAASKSGCGFPGSATVHLEOGGKTLVKDLRPGRDVLAAADGGRLYSDFL 240  
Db 181 HHCSSVAENSVAASKSGCGFPGSATVHLEOGGKTLVKDLRPGRDVLAAADGGRLYSDFL 240  
QY 241 TELDRDEGAKKVFYVIELEPERRELLTAAHLTFVAPHNDSGPTPGPSALFASVRPGQR 300  
Db 241 TELDRDEGAKKVFYVIELEPERRELLTAAHLTFVAPHNDSGPTPGPSALFASVRPGQR 300  
QY 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAHGTILINNVLASCAVIEHSHMAH 360  
Db 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAHGTILINNVLASCAVIEHSHMAH 360  
QY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPYAGIHWYSQLLYHIGTW 420  
Db 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPYAGIHWYSQLLYHIGTW 420  
QY 421 LLDSETHMPLGMAVKSS 437  
Db 421 LLDSETHMPLGMAVKSS 437

## RESULT 14

US-09-969-520A-4  
; Sequence 4, Application US/09969520A  
; Patent No. US20020177163A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: BEACHY, Philip A.  
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED  
; FILE REFERENCE: JH01670-1  
; CURRENT APPLICATION NUMBER: US/09/969,520A  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 60/235,153  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Altered sonic hedgehog protein  
US-09-969-520A-4

Query Match 99.2% Score 2270; DB 9; Length 437;

Best Local Similarity 99.3% Pred. No. 6.2e-196;

Matches 434; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLLLLARCFVILASLLVCPGLACGPGRGKRRHKKLTPLAYKOFIPVAEKTIGAS 60  
Db 1 MLLLLARCFVILASLLVCPGLACGPGRGKRRHKKLTPLAYKOFIPVAEKTIGAS 60  
QY 61 GRYEGKITRNSERKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALASVMNOMP 120  
Db 61 GRYEGKITRNSERKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALASVMNOMP 120  
QY 121 VRLAVTEGMDDEGHHSESLHYEGRAVDITTSDDRKYKGLAFLAAYEAGFDWYYYESKA 180  
Db 121 VRLAVTEGMDDEGHHSESLHYEGRAVDITTSDDRKYKGLAFLAAYEAGFDWYYYESKA 180  
QY 181 HHCSSVAENSVAASKSGCGFPGSATVHLEOGGKTLVKDLRPGRDVLAAADGGRLYSDFL 240  
Db 181 HHCSSVAENSVAASKSGCGFPGSATVHLEOGGKTLVKDLRPGRDVLAAADGGRLYSDFL 240  
QY 241 TELDRDEGAKKVFYVIELEPERRELLTAAHLTFVAPHNDSGPTPGPSALFASVRPGQR 300  
Db 241 TELDRDEGAKKVFYVIELEPERRELLTAAHLTFVAPHNDSGPTPGPSALFASVRPGQR 300  
QY 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAHGTILINNVLASCAVIEHSHMAH 360  
Db 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAHGTILINNVLASCAVIEHSHMAH 360

QY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPYAGIHWYSQLLYHIGTW 420  
Db 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPYAGIHWYSQLLYHIGTW 420  
QY 421 LLDSETHMPLGMAVKSS 437  
Db 421 LLDSETHMPLGMAVKSS 437

## RESULT 15

US-09-969-520A-8  
; Sequence 8, Application US/09969520A  
; Patent No. US20020177163A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: BEACHY, Philip A.  
; TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED  
; FILE REFERENCE: JH01670-1  
; CURRENT APPLICATION NUMBER: US/09/969,520A  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: US 60/235,153  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Altered sonic hedgehog protein  
US-09-969-520A-8

Query Match 99.0% Score 2266; DB 9; Length 437;

Best Local Similarity 99.1% Pred. No. 1.4e-195;

Matches 433; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLLLLARCFVILASLLVCPGLACGPGRGKRRHKKLTPLAYKOFIPVAEKTIGAS 60  
Db 1 MLLLLARCFVILASLLVCPGLACGPGRGKRRHKKLTPLAYKOFIPVAEKTIGAS 60  
QY 61 GRYEGKITRNSERKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALASVMNOMP 120  
Db 61 GRYEGKITRNSERKELTPNYPNDIIFKDEENTGADRLMTORCKDKLNALASVMNOMP 120  
QY 121 VRLAVTEGMDDEGHHSESLHYEGRAVDITTSDDRKYKGLAFLAAYEAGFDWYYYESKA 180  
Db 121 VRLAVTEGMDDEGHHSESLHYEGRAVDITTSDDRKYKGLAFLAAYEAGFDWYYYESKA 180  
QY 181 HHCSSVAENSVAASKSGCGFPGSATVHLEOGGKTLVKDLRPGRDVLAAADGGRLYSDFL 240  
Db 181 HHCSSVAENSVAASKSGCGFPGSATVHLEOGGKTLVKDLRPGRDVLAAADGGRLYSDFL 240  
QY 241 TELDRDEGAKKVFYVIELEPERRELLTAAHLTFVAPHNDSGPTPGPSALFASVRPGQR 300  
Db 241 TELDRDEGAKKVFYVIELEPERRELLTAAHLTFVAPHNDSGPTPGPSALFASVRPGQR 300  
QY 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAHGTILINNVLASCAVIEHSHMAH 360  
Db 301 VYVAERGGDRLLPAAVHSTLREERAGAYAPLTAHGTILINNVLASCAVIEHSHMAH 360  
QY 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPYAGIHWYSQLLYHIGTW 420  
Db 361 RAFAFRLAHLAALAPARTDGGGGSIPAAOSATEARGAEPYAGIHWYSQLLYHIGTW 420  
QY 421 LLDSETHMPLGMAVKSS 437  
Db 421 LLDSETHMPLGMAVKSS 437

Search completed: February 20, 2003, 10:14:42  
Job time : 8.8443 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:06:28 ; Search time 28.1092 Seconds  
(without alignments)  
3203.314 Million cell updates/sec

Title: US-09-827-110A-13  
Perfect score: 2289  
Sequence: 1 MLLILARCLVLIASSLVC.....GTWLDSETHMPLGMVKKSS 437

Scoring table: BLAST62  
Gap 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SPREMBL\_21:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mmc:\*  
9: sp-organella:\*  
10: sp-phage:\*  
11: sp-plant:\*  
12: sp-rodent:\*  
13: sp-virus:\*  
14: sp-vertebrate:\*  
15: sp-unclassified:\*  
16: sp\_rvirus:\*  
17: sp\_bacteriophage:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1537.5	67.2	434	13	057404 pleurodeles
2	1451	63.4	414	13	09W709 paralichth
3	1326	57.9	406	13	057567 Q9U526 branchiost
4	1132.5	49.5	415	5	09U526 branchiost
5	1124.5	49.1	415	5	017499 branchiost
6	1063.5	46.5	442	13	073803 fugu rubrip
7	1040.5	45.5	447	5	09GRA8 O9G8A8 gryllus bim
8	915.5	40.0	410	5	061676 O61676 lytechinus
9	842	36.8	161	11	09R179 O9R179 rattus norv
10	736	32.2	150	13	09YGV7 O9YGV7 ambystoma m
11	732	32.0	139	6	09XSI6 O9XSI6 bos taurus
12	710	31.0	138	13	09W6C1 O9W6C1 eleutheroda
13	698	30.5	177	11	09WZ29 O9WZ29 rattus norv
14	628	27.4	185	5	096699 O96699 junonia coe
15	528	23.1	99	13	08G6N4 O8G6N4 scyllorhinu
16	524	22.9	119	13	042128 O42128 oryzias lat

17	518	22.6	129	11	09WUP6 O9WUP6 rattus norv
18	489	21.4	137	13	042234 O42234 coturnix co
19	395	17.3	87	5	09TX30 O9TX30 anopheles g
20	389	17.0	80	13	042441 O42441 oryzias lat
21	339	14.8	63	13	09T814 O9T814 rana catesb
22	329	14.4	64	13	09PRF5 O9PRF5 oryzias lat
23	295	12.9	60	5	09S2D8 O9S2D8 hydra atten
24	289	12.6	56	13	09O2R0 O9O2R0 oryzias lat
25	234	10.2	49	5	09TX33 O9TX33 hludo medl
26	232	10.1	49	5	09TX31 O9TX31 trilobium c
27	187.5	8.2	125	11	09ESH3 O9ESH3 rattus norv
28	175	7.6	557	5	094410 O94410 caenorhabd1
29	175	7.6	557	5	094129 O94129 caenorhabd1
30	174	7.6	54	13	042233 O42233 coturnix co
31	172.5	7.5	48	5	09TX32 O9TX32 stryomon co
32	170	7.4	615	5	091573 O91573 caenorhabd1
33	153.5	6.7	5	5	021535 O21535 caenorhabd1
34	148	6.5	790	5	022872 O22872 caenorhabd1
35	148	6.5	1203	5	021835 O21835 caenorhabd1
36	146	6.4	481	5	045992 O45992 caenorhabd1
37	140.5	6.1	550	5	045273 O45273 caenorhabd1
38	140.5	6.1	550	5	094130 O94130 caenorhabd1
39	130.5	5.7	868	5	09XV14 O9XV14 caenorhabd1
40	128	5.6	1021	5	09XUV2 O9XUV2 caenorhabd1
41	127	5.5	485	5	094128 O94128 caenorhabd1
42	118	5.2	598	2	083032 O83032 streptomyce
43	117	5.1	1407	16	069826 O69826 streptomyce
44	113	4.9	599	16	09L010 O9L010 streptomyce
45	106	4.6	969	16	092RRO O92RRO rhizobium m

#### ALIGNMENTS

RESULT 1  
ID 057404 PRELIMINARY; PRT; 434 AA.  
AC 057404;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Sonic hedgehog-related protein.  
GN pw-SH.  
OS Pleurodeles waltl11 (Iberian ribbed newt).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae;  
OC Pleurodeles  
CX NCBI\_TaxID=8319;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Caubit X., Nicolas S., Le Parco Y.;  
RT "Pleurodeles sonic hedgehog".  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF003532; AAB94412.1; -.  
DR HSSP; Q62226; LVH.  
DR MEROPS; C46.002; -.  
DR InterPro: IPR001767; Hedgehog\_hnt.  
DR InterPro: IPR003586; Hedgehog\_hntC.  
DR InterPro: IPR003587; Hedgehog\_hntN.  
DR InterPro: IPR003203; Intein.  
DR InterPro: IPR003203; Intein.  
DR Pfam: PF01085; HH\_signal; 1.  
DR Pfam: PF01085; HH\_signal; 1.  
DR PRINTS; PF00632; SONICHHOG.  
DR PRODOM; PD003042; HH\_signal; 1.  
DR SMART; SM00305; HntC; 1.  
DR SMART; SM00306; HntN; 1.  
SQ SEQUENCE 434 AA; 48421 MW; A9495E367151AE74 CRC64;

Query Match 67.2%; Score 1537.5; DB 13; Length 434;  
Best Local Similarity 69.5%; Pred. No. 1.8e-115;  
Matches 310; Conservative 38; Mismatches 73; Indels 25; Gaps 6;

OY	2	LLLLACPLVLIASLLACPLACGPGGFCRRHPPKPLPLAYKOPIPNVAETKTGASG	61
	:	:                   :	
Db	4	MILLARDDLGIFISLILVPLGCGGPGRGIGQRRRPOKILPLAYKOPIPNVEKTTGASG	63
OY	62	RYECKITRNSERPEKELPPNPNDDIIFFKDEENTGADRLMTORCKDKLALAI SVANOMPGV	121
	:		
Db	64	RYEVKITRNSERPOELPPNPNDDIIFFKVEENTGADRLMTORCKDKLNSIAI SVANOMPGV	123
OY	122	RLRTEGDEGCHHSEESLHAYEGRAVDITTSDDPRSRKYGMLARLAVNAGDMWYEEKAH	181
	:		
Db	124	KLRTGEGDEGCHHSEESLHAYEGRAVDITTSDDPRSRKYGMLARLAVNAGDMWYEEKAH	183
OY	182	IHCYSKAENSAVAAKSGGCFPCSAFVHLEOGGCTKVLKPLRPEDRYLADDDGGRLLYSDFLT	241
	:		
Db	184	IHCYSKAENSAVAAVSGGCFPCSAFVHLEOGGRIPYVKGLKPNRRLAVNADVEGRLLYSDFLT	243
OY	242	FLUDRECAKKFYVYIETLEPRERLLTLAAHLFVAAPHNDGSPYGP-----SALF	291
	:		
Db	244	FMDEETLARKFYVYIETSLPRERLRLTLAAHLFVAPOCHPGNASGANDRSGKFCGHRRPFSMF	303
OY	292	ASRYAPGGRVYVVAERGGDRLLTAAVHASYTLREEGAAYAPLTAHGHTILINRYLASCYA	351
	:		
Db	304	ASSVAPGRV- LTEDREG- KGLRATVYDRYL- DEANGATAVPATVATGTVIDRYLASCYA	360
	:		
OY	352	VIEEHSMWARFAFPRLLAHLAHLAIPARTDGGGGGSIIPAAQSAATERGAEPTAGIHHMS	411
	:		
Db	361	VIEEHSMWARFAFPRLVGFGILTSFSPQDYSMS-----PRAPSGAE-----GVHMS	408
OY	412	QLLVHICWLLDSEFMHPLGAAVSS	437
	:		
Db	409	EILYRIGTWVLAQADTIHPLGMAASS	434

RESULT 2			
09W709	PRELIMINARY;	PRT;	414 AA.
ID 09W709			
AC 09W709;			
DT 01-NOV-1999 (TREMBREL, 12, Created)			
DT 01-NOV-1999 (TREMBREL, 12, Last sequence update)			
DT 01-MAR-2002 (TREMBREL, 20, Last annotation update)			
DE Sonic hedgehog.			
GN SHH.			
OS Paralicthys olivaceus (Flounder).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC Acanthomorphi; Acanthopterygii; Percomorpha; Pleuronectiformes;			
OC Pleuronectidae; Paralicthyidae; Paralicthys.			
OX NCBI_Taxid=8255;			
RA [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE-99238226; PubMed-10223710;			
RA Suzuki T., Ichiro O., Kurokawa T.;			
"Retinoic acid given at late embryonic stage depresses sonic hedgehog			
RT and Hoxd-4 expression in the pharyngeal area and induces skeletal			
RT malformation in flounder (Paralicthys olivaceus) embryos.";			
RL Dev. Growth Differ. 41:143-152(1999).			
DR EMBL: AB029748; BAA82360.1; -			
DR HSSP: Q62226; 1VHN.			
DR MEMOPS: C46_001; -			
DR InterPro: IPR001767; Hedgehog_hint.			
DR InterPro: IPR003586; Hedgehog_hintc.			
DR InterPro: IPR003587; Hedgehog_hintn.			
DR InterPro: IPR003020; HH_signal.			
DR InterPro: IPR002203; Intein.			
DR InterPro: IPR001657; SonicHH.			
DR Pfam: PF01085; HH_signal; 1.			
DR Pfam: PF01079; Hintc; 1.			
DR PRINTS: PR00632; SONICHHOG.			
DR ProDom: PD003042; HH_signal; 1.			
DR SMART: SM00305; Hintc; 1.			
DR SMART: SM00306; Hintn; 1.			
SO SEQUENCE 414 AA; 45945 MW; 50607BF3DB7CDAA3 CRC64;			

Query Match	63.48;	Score 1451;	DB 13;	Length 414;
Best Local Similarity	65.88;	Pred. No. 1.5e-108;		
Matches 288;	Conservative 53;	Mismatches 71;	Indels 26;	Gaps 7

[illegible]

RESULT 3	057567	PRELIMINARY;	PRT;	406 AA.
ID	057567			
AC	057567;			
DT	01-JUN-1998 (TrEMBLrel. 06, Created)			
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Hedgehog segment polarity homolog.			
OS	Neurospora crassa (Fungal Fungi)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; I			
OC	Amphibia; Batrachia; Caudata; Salamandroides; Salamand			
OC	Neurospora crassa.			
OX	NCBI_TaxID=8316;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Stark D.R., Gates P.B., Brookes J.P., Ferretti P.;			
RT	"Hedgehog homologue from <i>Neurospora crassa</i> viridescens.",			
RL	Dev. Dyn. 0:0-(1998).			
CC	-1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE			
CC	PHOSPHORIBOSYLTRANSFERASE FAMILY.			
CC	EMBL: AF047466; AAC03108.1; -.			
DR	HSPB; 062226; IVH.			
DR	InterPro: IPR0011757; Hedgehog_hint.			
DR	InterPro: IPR003586; Hedgehog_hintC.			
DR	InterPro: IPR003587; Hedgehog_hintN.			
DR	InterPro: IPR000320; HH_signal.			
DR	InterPro: IPR002375; Pr/PY_rp_transf.			
DR	InterPro: IPR001657; SonicHH.			
DR	Pfam: PF01085; HH_signal; 1.			
DR	Pfam: PF01079; Hint; 1.			
DR	PRINTS: PR00632; SONICHHOG.			
DR	Prodom: PD003042; HH_signal; 1.			
DR	SMART: SM00305; Hintn; 1.			
DR	SMART: SM00306; Hintn; 1.			
DR	PROSITE: PS00103; PUR_PYR_TRANSFER; 1.			







```

DB 121 TWHLEGGTKLVKDLSPGDRVLAADQGRILYSDFLTFLDR 161
|||||

RESULT 10
ID 09YGV7 PRELIMINARY; PRT; 150 AA.
AC 09YGV7:
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Ambystoma mexicanum (Axolotl).
OC Eukaryota; Metazoa; Chordata; Salamandroidae; Ambystomatidae;
OC Amphibia; Batrachia; Caudata;
OC Ambystoma.
OX NCBI_TaxID=8296;
RN [1]
RP SEQUENCE FROM N.A.
RA Torok M.A., Ispizua-Belmonte J.C., Gardiner D.M., Bryant S.V.;
RL Submitted (Oct-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031480; AAD18128.1; -.
DR HSSP; Q62226; 1VHH.
DR InterPro; IPR001767; Hedgehog_hint.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; SonicHH.
DR Pfam; PF01085; HH_signal; 1.
DR Pfam; PF01079; Hint; 1.
DR PRINTS; PR00632; SONICHHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 150
SQ SEQUENCE 150 AA; 16599 MW; 986F65037A69A8AC CRC64;

Query Match
Best Local Similarity 32.2%; Score 736; DB 13; Length 150;
Matches 140; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 84 DIFKDEENTGADRLMTORCKDKLNLAISVNMOPGVRLRYTEGDEDEHHSESLAYE 143
|||||
DB 1 DIFKDEENTGADRLMTORCKDKLNLAISVNMOPGVRLRYTEGDEDEHHSESLAYE 60
|||||

OY 144 GRAVDITTSDRDRSKYGMALARLAVEGFPWYVESKAHHCYKKAANSVAAKSGGCPFS 203
|||||
DB 61 GRAVDITTSDRDRSKYGMALARLAVEGFPWYVESKAHHCYKKAANSVAAKSGGCPFS 120
|||||

OY 204 ATYHLEGGTKLVKDLSPGDRVLAADQGR 233
|||||
DB 121 AKVTLEHGVTRPVKDLRPGRVLAADQGR 150
|||||

RESULT 11
O9XS16 PRELIMINARY; PRT; 139 AA.
ID 09XS16:
AC 09XS16:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue-Tooth Germ;
RC Koyama E., Iwamoto M., Ohmori T., Kurisu K., Wu C., Ookura T.,
RA Basbair M.M., Tucker T., Pacifici M.;
RT "Development of Stratum Intermedium and its Role as a Sonic Hedgehog-
RT Signaling Structure During Odontogenesis.";
RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
EMBL; AF144100; AAD33926.1; -.

```

```

DR HSSP; Q62226; 1VHH.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; SonicHH.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 139
SQ SEQUENCE 139 AA; 15961 MW; C83B0B96982046B1 CRC64;

Query Match
Best Local Similarity 32.0%; Score 732; DB 6; Length 139;
Matches 137; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 44 AVKQFIPNVAEKTIGASGRYEKRTNSRERKELTPYNDPIFKDEENTGADRLMTORC 103
|||||
DB 1 AVKQFIPNVAEKTIGASGRYEKRTNSRERKELTPYNDPIFKDEENTGADRLMTORC 60
|||||

OY 104 KDKLNLAISVNMOPGVRLRYTEGDEDEHHSESLAYEGRAVDITTSDRDRSKYGMALA 163
|||||
DB 61 KDKLNLAISVNMOPGVRLRYTEGDEDEHHSESLAYEGRAVDITTSDRDRSKYGMALA 120
|||||

OY 164 RLAVEAGFPWYVESKAH 182
|||||
DB 121 RLAVEAGFPWYVESKAH 139
|||||

RESULT 12
O9W6C1 PRELIMINARY; PRT; 138 AA.
ID 09W6C1:
AC 09W6C1:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Sonic hedgehog protein (Fragment).
OS SHH.
OC Eleutherodactylus coqui (Puerto Rican coqui).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;
OC Eleutherodactylus.
OX NCBI_TaxID=57060;
RN [1]
RP SEQUENCE FROM N.A.
RA Carl T.F., Richardson M.K., Olsson L., Schlosser G., Klimkowski M.W.,
RA Hanken J.;
RT "Differences in vertebrate limb development revealed by studies of the
RT direct developing frog E. coqui.";
RL Submitted (Dec-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF113403; AAD23436.1; -.
DR HSSP; Q62226; 1VHH.
DR InterPro; IPR000320; HH_signal.
DR InterPro; IPR001657; SonicHH.
DR Pfam; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR ProDom; PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15751 MW; FF4156A17F4681F0 CRC64;

Query Match
Best Local Similarity 31.0%; Score 710; DB 13; Length 138;
Matches 133; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 49 IPNVAEKTIGASGRYEKRTNSRERKELTPYNDPIFKDEENTGADRLMTORCKDKLN 108
|||||
DB 1 IPNVAEKTIGASGRYEKRTNSRERKELTPYNDPIFKDEENTGADRLMTORCKDKLN 60
|||||

OY 109 ALAISVNMOPGVRLRYTEGDEDEHHSESLAYEGRAVDITTSDRDRSKYGMALARLAVE 168
|||||
DB 61 ALAISVNMOPGVRLRYTEGDEDEHHSESLAYEGRAVDITTSDRDRSKYGMALARLAVE 120
|||||

OY 169 AGFDWYVESKAHHCYV 186
|||||

```

Db 121 AGEDWYTESKAHICSV 138

RESULT 13

09WV29 PRELIMINARY; PRT; 177 AA.

AC 09WV29; (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Indian hedgehog protein (Fragment).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;

RA Garges P.L., Meyer R.A., Jr., Brown C.A., Price D.K.;

RT "Indian hedgehog in rat."

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF162914; AAD45372.1; -

DR HSSP: Q62226; 1VHH.

DR MEROPS: C46.003; -

DR InterPro: IPR001767; Hedgehog\_hnt.

DR InterPro: IPR003587; Hedgehog\_hntn.

DR InterPro: IPR00320; HH\_signal.

DR InterPro: IPR001657; SonichH.

DR Pfam: PF01085; HH\_signal; 1.

DR Pfam: PF01079; Hntc; 1.

DR PRINTS: PR00632; SONICHOG.

DR PRODOM: PD003042; HH\_signal; 1.

DR SMART: SM00306; Hntn; 1.

FT NON\_TER 1

FT NON\_TER 1

SQ SEQUENCE 177 AA; 19739 MW; CAS871626A565FE5 CRC64;

Query Match 30.5%; Score 698; DB 11; Length 177;

Best Local Similarity 73.4%; Pred. No. 1.9e-48;

Matches 130; Conservative 17; Mismatches 30; Indels 0; Gaps 0;

QY 98 LMTORCKDKLNALAIYVNMOPGRLVTEGMDDEGHSESLAYEGRAVDITTSDDRS 157

Db 1 LMTORCKDKLNALAIYVNMOPGRLVTEGMDDEGHSESLAYEGRAVDITTSDDRS 60

QY 158 KYGMALAVAGFDWYTESKAHICSVKAKNSVAKSGCGPFSATVHLEOGCTKLK 217

Db 61 KYGLALAVAGFDWYTESKAHICSVKSEHSAKKTGCGFAGAOVHLETGERVALS 120

QY 218 DLREGDRLAADGGRLLYSDFLFDREDEGAKKVFYVETLEPRERLLLTAAHLF 274

Db 121 AVKPGDRLVAGDEGNPFESDVLFLDREPRRLRAFGVETDPPRRLLTAAHLF 177

RESULT 14

096699 PRELIMINARY; PRT; 185 AA.

AC 096699; (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE Hedgehog protein (Fragment).

GN HH.

OS Junonia coenia (Peacock butterfly) (Precis coenia).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Plekoptera; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditylata;

OC Papilionidae; Nymphalidae; Nymphalinae; Junonia.

OX NCBI\_TaxID=39708;

RN [1]

RP SEQUENCE FROM N.A.

RA Keys D.N., Lewis J., Selegue J.E., Pearson B.J., Goodrich L.V.,

RA Johnson R.L., Gates J., Scott M.P., Carroll S.B.;

RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot

RT evolution.";

RL Science 0:0-0(1999).

DR EMBL: AF117742; AAD08931.1; -

DR HSSP: Q62226; 1VHH.

DR InterPro: IPR00320; HH\_signal.

DR InterPro: IPR001657; SonichH.

DR Pfam: PF01085; HH\_signal; 1.

DR PRINTS: PR00632; SONICHOG.

DR PRODOM: PD003042; HH\_signal; 1.

FT NON\_TER 1

FT NON\_TER 1

SQ SEQUENCE 185 AA; 20745 MW; E9B9F4CE2FAAE662 CRC64;

Query Match 27.4%; Score 628; DB 5; Length 185;

Best Local Similarity 97.4%; Pred. No. 8.9e-43;

Matches 118; Conservative 20; Mismatches 30; Indels 2; Gaps 2;

QY 59 ASGRTEKTRNSERFELTPNPNPDILFEDENTGADRLMTORCKDKLNALAIYVNM 118

Db 3 ASGPPEGRTFDEKFFDLVNPYNPDIDFDEGTGADRLMTORCKDKLNALAIYVNM 62

QY 119 PGVRLVTEGMDDEGHSESLAYEGRAVDITTSDDRSKYGMALAVAGFPWYTES 178

Db 63 PGVRLVTEGMDDEGHSESLAYEGRAVDITTSDDRSKYGMALAVAGFPWYTES 122

QY 179 KAHICSVKAKNSVAKSGCGPFSATVHLEOGCTKLKDLRPGDRLVLA 228

Db 123 RSYIHCSVKTESVGT-GAGCFPGAVVHTENGFXD-IASLKGKQVLA 170

RESULT 15

080GN4 PRELIMINARY; PRT; 99 AA.

AC 080GN4; (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DE Sonic hedgehog (Fragment).

OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;

OC Elasmobranchii; Galeomorphi; Galeidae; Carcharhiniformes;

OC Scyllorhinidae; Scyllorhinus.

OX NCBI\_TaxID=7830;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21930458; Pubmed=11932743;

RA Tanaka M., Munsterberg A., Anderson W.G., Prescott A.R., Hazon N.,

RA Tickle C.;

RT "Fin development in a cartilaginous fish and the origin of vertebrate

RL Limbs.";

RL Nature 416:527-531(2002).

DR EMBL: AF393835; AAM08228.1; -

FT NON\_TER 1

FT NON\_TER 1

SQ SEQUENCE 99 AA; 11468 MW; 831864BDB34A1A20 CRC64;

Query Match 23.1%; Score 528; DB 13; Length 99;

Best Local Similarity 97.0%; Pred. No. 4e-35;

Matches 96; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 79 PNYPNPDILFEDENTGADRLMTORCKDKLNALAIYVNMOPGRLVTEGMDDEGHSE 138

Db 1 PNYPNPDILFEDENTGADRLMTORCKDKLNALAIYVNMOPGRLVTEGMDDEGHSE 60

QY 139 SLHTEGRAVDITTSDDRSKYGMALAVAGFPWYTES 177

Db 61 SLHTEGRAVDITTSDDRTKGMALAVAGFPWYTES 99

Search completed: February 20, 2003, 10:11:51

Job time : 29.1092 secs

**THIS PAGE BLANK (USPTO)**





PI Ingham PW, McMahon AP, Tablin CJ;  
 XX WPI, 1995-255060/33.  
 DR N-PSDB; AAQ91639.  
 XX  
 PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful  
 PT to treat degenerative nervous system disorder(s) and in gene  
 PT therapy.  
 XX  
 PS Claim 17; Page 143-45; 210pp; English.  
 XX  
 CC The sequence represents a human sonic hedgehog protein, homologous  
 CC to a Drosophila hedgehog protein (AA073357), and is encoded by a cDNA  
 CC isolated from a human fetal lung cDNA library. Probes and primers  
 CC derived from the sonic hedgehog gene may be used as diagnostic  
 CC agents for neuromuscular, autonomic or central nervous system  
 CC disorders, and the gene may also be used in gene therapy.  
 CC Antibodies generated from the protein may be used as therapeutic or  
 CC research reagents.  
 CC  
 SQ Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 16; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHPRKLTPLAYKQFIPNVAEKTIGASG 60  
 DB 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHPRKLTPLAYKQFIPNVAEKTIGASG 60  
 QY 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOMPCV 120  
 DB 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOMPCV 120  
 QY 121 KLRVTEGMEDGHSEESIHYEGRAVDITTSDRSKYGMRLAIVEAGFDWYTESKAH 180  
 DB 121 KLRVTEGMEDGHSEESIHYEGRAVDITTSDRSKYGMRLAIVEAGFDWYTESKAH 180  
 QY 181 IHCSVAENSVAKSGCGPGSATVHLEGGGTRKLVKDLSPGDRVLAADOGRLYSDFLT 240  
 DB 181 IHCSVAENSVAKSGCGPGSATVHLEGGGTRKLVKDLSPGDRVLAADOGRLYSDFLT 240  
 QY 241 FLDRODGAKKVYVITRPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGALG 300  
 DB 241 FLDRODGAKKVYVITRPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGALG 300  
 QY 301 PALFASRRVPGORVYVAERDGRRLPAVAHSVTLSEEAAGAVPLTAOGTILINRYL 360  
 DB 301 PALFASRRVPGORVYVAERDGRRLPAVAHSVTLSEEAAGAVPLTAOGTILINRYL 360  
 QY 361 ASCYAVIEEHSMAHRAFAFRLAHLLALAPARTDRGDSGGGGRGGRVALTAPGA 420  
 DB 361 ASCYAVIEEHSMAHRAFAFRLAHLLALAPARTDRGDSGGGGRGGRVALTAPGA 420  
 QY 421 AARPAGAGATGHTWYSQLLYOIGTWLIDSEALHPGLGMANKSSXSRACGAGAGA 475  
 DB 421 ADAPGAGATGHTWYSQLLYOIGTWLIDSEALHPGLGMANKSSXSRACGAGAGA 475

## RESULT 2

AA05859  
 ID AA05859 standard; Protein; 475 AA.

AC AA05859;

DT 02-AUG-1999 (first entry)

DE Human Sonic hedgehog Shh protein.

XX Sonic hedgehog; Shh; human; epithelial tissue; epithelium;  
 XX cutaneous tissue; skin; hair; wound healing; vulvectomy;  
 KM burn; skin grafting; pressure sore; ulcer; ulcerative colitis;  
 KM alopecia; psoriasis; keratosis; acne; comedogenic lesion;

KW folliculitis; pseudo-folliculitis; keratoacanthoma; callusities;  
 KW Darier's disease; scar; autoimmune disease; pemphigus;  
 KM epidermolysis; lupus lesion; desquamative lesion; carcinoma;  
 XX therapy; hedgehog therapeutic; ptc therapeutic; patched.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 463 /note="encoded by NNN"

XX W09920298-A1.

XX 29-APR-1999.

XX 20-OCT-1998; 98MO-0522227.

XX 11-SEP-1998; 98US-0151999.

XX 20-OCT-1997; 97US-0955552.

XX (ONTO-) ONTOGENY INC.

XX Wang EA;

XX WPI, 1999-288170/24.

XX N-PSDB; AA05622.

XX Claim 26; Page 127-128; 146pp; English.

The present sequence represents human Sonic hedgehog protein Shh. The invention relates to a method for modulating the growth state an epithelial cell by ectopically contacting the epithelial cell, in vitro or in vivo, with a hedgehog therapeutic (i.e. a hedgehog polypeptide or gene therapy construct) or ptc therapeutic (i.e. a small organic molecule that mimics the effect of hedgehog proteins on patched signalling, or activates or potentiates patched signalling) in an amount effective to alter the rate of proliferation of the epithelial cell. The hedgehog therapeutic preferably comprises at least a bioactive extracellular portion of a hedgehog protein (see AA05854-62) encoded by a vertebrate hedgehog gene (see AA05617-25), especially a human hedgehog gene. Promotion of proliferation of epithelial cells can be used to control a wound healing process in e.g. burn treatment, skin regeneration, skin grafting, pressure sore treatment, dermal ulcer treatment, post surgery scar reduction or treatment of ulcerative colitis (claimed). It can also be used to induce hair growth for the treatment of alopecia (claimed). Inhibition of the growth of epithelial tissue can be used to treat or prevent hyperplastic or neoplastic conditions, e.g. psoriasis, keratosis, acne, comedogenic lesions, folliculitis and pseudo-folliculitis, keratoacanthoma, callusities, Darier's disease, keloids, hypertrophic scars, or autoimmune disorders, e.g. aphthous ulcers, pemphigus vulgaris, pemphigus foliaceus, pemphigus vegetans, pemphigus erythematous, epidermolysis, lupus lesions, desquamative lesions or carcinomas. The methods can also be used to counteract the effects of ageing on skin.

SQ Sequence 475 AA;

Query Match 99.9%; Score 2467; DB 20; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHPRKLTPLAYKQFIPNVAEKTIGASG 60  
 DB 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHPRKLTPLAYKQFIPNVAEKTIGASG 60  
 QY 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOMPCV 120  
 DB 61 RREGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNOMPCV 120

Qy	121	KLNTJTEGMDDEGHHSSESLHTEGRAVDTITSDRDRSKYGMRLARLAVAGDWMYYESKAH	180
Db	121	KLRLTEGMDDEGHHSSESLHTEGRAVDTITSDRDRSKYGMRLARLAVAGDWMYYESKAH	180
Qy	181	IHC5VKAEN5VAASGGCFPGSATAVHLEOGGTKLVKQLSPGDRVLADDDGRLTYSFLT	240
Db	181	IHC5SKAEN5VAASGGCFPGSATAVHLEOGGTKLVKQLSPGDRVLADDDGRLTYSFLT	240
Qy	241	FLDRDDGAKKFFYYIETREPERRLTLTAHLTLFVAPHNDSATGCEPASGSPGSGALG	300
Db	241	FLDRDDGAKKFFYYIETREPERRLTLTAHLTLFAPHNDSATGCEPASGSPGSGALG	300
Qy	301	PRALFASRVRRGQRYVVAERDGRRLTLPAAVHSVTLSSEAAGAYAPLTAOGTILINRVL	360
Db	301	PRALFASRVRRGQRYVVAERDGRRLTLPAVHSVTLSSEAAGAYAPLTAOGTILINRVL	360
Qy	361	ASCAVIEEHSWMAHRAFPFLAHALLAALAPARTDRGGDSGGDRGGGGGAYALTPGA	420
Db	361	ASCAVIEEHSWMAHRAFPFLAHALLAALAPARTDRGGDSGGDRGGGGGAYALTPGA	420
Qy	421	ADAPGAGTAAIHMYSQLLYIGTWLIDSEALHPLGMAVVS55XSRGAGGAGAREGA	475
Db	421	ADAPGAGTAAIHMYSQLLYIGTWLIDSEALHPLGMAVVS55XSRGAGGAGAREGA	475

CC Also claimed is a method for treatment or prevention of disorders  
CC of, or surgical or cosmetic repair of, such muscle tissues, by  
CC administering a hedgehog polypeptide or ptc therapeutic. The  
CC disorder may be muscle atrophy, in particular skeletal muscle  
CC atrophy or cardiac muscle atrophy, cachexia, or muscular myopathy  
CC (all claimed). The hedgehog polypeptide or ptc therapeutic can  
CC inhibit growth of myoblastic-derived tissue to provide treatment of  
CC hyperplastic or neoplastic growth of muscle tissue such as in  
CC myoblastic sarcoma (also claimed). The hedgehog therapeutic  
CC preferably comprises at least a bioactive extracellular portion of  
CC a hedgehog protein (see AA0510-19) encoded by a vertebrate hedgehog  
CC gene (see AAX25098-107), especially a human hedgehog gene.  
XX  
XX  
SQ Sequence 4/75 AA;

XX	FN		MO3904775-A2.	
XX	PD		04-FEB-1999.	
XX	PE		24-JUL-1998;	98WO-US15419.
XX	PR		24-JUL-1997;	97US-0900220.
XX	PA		(ONTO-) ONTOGENY INC.	
XX	PI	Mahantappa NK,	Miso N,	Pang K,
XX	DR	WPI; 1999-142578/12.		
XX	DR	N-PSDB; MAX07276.		
PT	PT	Increasing the survival of neuronal, dopaminergic and GABA-nergic		
PT	PT	cells - by using a ptc therapeutic such as a protein kinase		
PT	PT	inhibitor, or an agent derived from hedgehog polypeptides, useful in		
PT	PT	the treatment of Parkinson's disease		
PS	PS	Disclosure; Page 93-95; 138pp; English.		
CC	XX			
CC	CC	This polypeptide is human Shh Sonic hedgehog protein. The		
CC	CC	invention is based on the finding that hedgehog proteins are useful		
CC	CC	as protective agents in the treatment and prophylaxis of		
CC	CC	neurodegenerative disorders resulting from the loss of dopaminergic		
CC	CC	and/or GABA-nergic neurons, or the general loss of tissue from the		
CC	CC	substantia nigra. Exemplary disorders include Parkinson's disease,		
CC	CC	Huntington's disease (both claimed), amyotrophic lateral sclerosis		
CC	CC	and cerebral ischemia. The invention relates to hedgehog		
CC	CC	therapeutics (i.e. hedgehog polypeptides and gene therapy constructs		
CC	CC	e.g. constructs encoding recombinant hedgehog polypeptides and		
CC	CC	trans-activation constructs for altering hedgehog gene regulatory		
CC	CC	sequences) and ptc therapeutics (i.e. agents which mimic the effect		
CC	CC	of naturally occurring hedgehog proteins on patched signaling)		
CC	CC	that are effective in both human and animal subjects. Human Ihh		
CC	CC	and Dhh polypeptides (see AAW97763-64) are preferred. The products		
CC	CC	can also be used for the maintenance of differentiated neurons in		
CC	CC	cultures, and to enhance the implantation of such neuronal cells in		
CC	CC	an animal. They can be used to prevent or treat neurodegenerative		
CC	CC	conditions arising from the use of certain drugs, and in the		
CC	CC	prevention and/or treatment of hypoxia, e.g. as a neuroprotective		
CC	CC	agent.		
SO	XX	Sequence	475 AA;	
		Query Match	99.9%; Score 2467;	DB 20; Length 475;
		Best Local Similarity	100.0%; Pred. No. 5.9e-218;	
		Matches 475; Conservative	0; Mismatches	0; Indels
				Gaps
OY		1	MLLARCLLVVSSLLVCGLACPGRGFGRRRPPKLTPLAYKQFIPNVAEKTLLGASG	60
DB		1	MLLARCLLVVSSLLVCGLACPGRGFGRRRPPKLTPLAYKQFIPNVAEKTLLGASG	60
OY		61	RYEGRISNRSEFKELTPNYNPDIIFKDEENTGADRLMTORCKDKLMLAISVMNQPGV	120
DB		61	RYEGKISRNSERFKELTPNYNPDIIFKDEENTGADRLMTORCKDKLMLAISVMNQPGV	120
OY		121	KLRVTGEWDGDSHSSESLHYEGRAVDITTSRDRSKTKGMARLAVEAGFDVVYYESKAH	180
DB		121	KLRVTGEWDGDSHSSESLHYEGRAVDITTSRDRSKTKGMARLAVEAGFDVVYYESKAH	180
OY		181	IHCSTAKSNSVAASKGGCFPGSATYHLBOGCGTLVKDLSPGRVLAADDQGSLTSDFLT	240
DB		181	IHCSTAKSNSVAASKGGCFPGSATYHLBOGCGTKLYKDLSPGRVLAADDQGSLTSDFLT	240
OY		241	FLDRDDGAKKYFYVETREPRERLLTLAHLLETVAPHNDSATGEPEASSGSPPSGALG	300
DB		241	FLDRDDGAKKYFYVETREPRERLLTLAHLLETVAPHNDSATGEPEASSGSPPSGALG	300
OY		301	PRALPASVRPQRYYVVAERDGRRLLPAAVHSYTLSEAGAAPAFLTAOSTITILNRYL	360
DB		301	PRALPASVRPQRYYVVAERDGRRLLPAAVHSYTLSEAGAAPAFLTAOSTITILNRYL	360

QY	361	ASCAVIEHSHMARAFAPFLIALAALAPARTDGGSGGDRGGGGRVALTPGA	420
QY	361	ASCAVIEHSHMARAFAPFLIALAALAPARTDGGSGGDRGGGGRVALTPGA	420
Db	361	ASCAVIEHSHMARAFAPFLIALAALAPARTDGGSGGDRGGGGRVALTPGA	420
QY	421	ADAPGAGTAGIHWYSOLLYOIGTWLDSALHPGMAVKSXSRGAGGARGGA	475
Db	421	ADAPGAGTAGIHWYSOLLYOIGTWLDSALHPGMAVKSXSRGAGGARGGA	475
RESULT 5			
ID	AAW94473		
AC	AAW94473	standard; Protein; 475 AA.	
XX	AAW94473;		
XX			
DT	29-APR-1999	(first entry)	
DE		Human Shh hedgehog protein sequence.	
XX			
KW		Patched; hedgehog; ptc therapeutic; neuroprotective; neuronal cell;	
KW		brain infarction; cerebral infarction; transient ischaemic attack;	
KW		stroke; cerebral infarct volume; spinal cord; oedema; trauma;	
KW		haemorrhage; encephalomyelitis; coronary bypass; cerebral hypoxia.	
OS		Homo sapiens.	
XX			
FT	Key	Location/Qualifiers	
FT	Misc-difference 463	/label= unknown	
FT		/note= "encoded by NNN"	
XX			
PN	WC9900117-A2.		
PD			
PD	07-JAN-1999.		
XX			
PF	26-JUN-1998;	98WC-US13387.	
XX			
PR	27-JUN-1997;	97US-0883656.	
XX			
PA	(ONTO-) ONTOGENY INC.		
XX			
P1	Mahanthappa NK;		
XX			
DR	WPI; 1999-095458/08.		
DR	N-PSDB; AAX16187.		
XX			
FT	Method for limiting damage to neurons caused by ischaemic or epoxic		
FT	conditions - is used for the treatment and prevention of e.g.		
PT	cerebral infarction, stroke and transient ischaemic attacks		
XX			
PS	Disclosure; Page 72-74; 104pp; English.		
XX			
CC	A method has been developed for limiting the damage to neuronal cells by		
CC	ischaemic or epoxic conditions by administering a ptc (patched)		
CC	therapeutic agent to reduce cerebral infarct volume (CIV). Damage to		
CC	neuronal cells can also be limited by administering a gene activation		
CC	construct which recombines with the genomic hedgehog gene to provide a		
CC	heterologous transcription regulator linked to the coding region of this		
CC	gene. Administration of the ptc therapeutic agent is used to protect		
CC	cerebral tissues against ischaemic injury; to treat cerebral infarct or		
CC	ischaemia, stroke (thrombotic or embolic) and transient ischaemic		
CC	attacks. It may also be used as a prophylactic in many other cases of		
CC	injury to the brain or spinal cord, oedema caused by trauma, haemorrhage		
CC	and encephalomyelitis, or in conjunction with (coronary bypass) surgery.		
CC	Treatment (which may be prophylactic) is used where ischaemic/epoxic		
CC	conditions may cause cerebral hypoxia, or progressive loss of neurons		
CC	due to oxygen depletion, including in patients with hypotension. The		
CC	treatment reduces CIV by at least 25, particularly at least 70%. The		
CC	present sequence represents a hedgehog sequence given in the present		
XX			
XX	Sequence	475 AA;	
XX			

Query Match 99.9%; Score 2467; DB 20; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACPGRCGRRPKKLPPLAVAKOPIPNVAEKTGLASG 60  
 D 1 MLLARCLLVVSSLLVCSGLACPGRCGRRPKKLPPLAVAKOPIPNVAEKTGLASG 60  
 QY 61 RYEGKISNSERFKELPNPNPDIIFFDEENTGADRLMTORCKKLNALAISSVMNQPGV 120  
 D 61 RYEGKISNSERFKELPNPNPDIIFFDEENTGADRLMTORCKKLNALAISSVMNQPGV 120  
 QY 121 KLRVTEGDEDDHSEESLHYEGRAVDITTSDRRSKYGMARLAVEGPDWYYESKAH 180  
 D 121 KLRVTEGDEDDHSEESLHYEGRAVDITTSDRRSKYGMARLAVEGPDWYYESKAH 180  
 QY 181 IHCSVKAENSVAAKSGCGFPGSATVHLEOGCTKLVDLSPEDRYLAADDGRLLYSDFLT 240  
 D 181 IHCSVKAENSVAAKSGCGFPGSATVHLEOGCTKLVDLSPEDRYLAADDGRLLYSDFLT 240  
 QY 241 FLDDDDGAKKVFYIETREPERELLTPAHLFVAPHNDSATGPEASSSGSPSGALG 300  
 D 241 FLDDDDGAKKVFYIETREPERELLTPAHLFVAPHNDSATGPEASSSGSPSGALG 300  
 QY 301 PRALFASVRPGORVYVAERDGRRLTPAAVHSVTLSEEAAGAYAPLTAGTLLINRVL 360  
 D 301 PRALFASVRPGORVYVAERDGRRLTPAAVHSVTLSEEAAGAYAPLTAGTLLINRVL 360  
 QY 361 ASCAIVAEHSHWARAFAPRLAHLAALAPARTDRGDSGGDGGGGRVALTAPGA 420  
 D 361 ASCAIVAEHSHWARAFAPRLAHLAALAPARTDRGDSGGDGGGGRVALTAPGA 420  
 QY 421 ADAGAGATAGIHMYSOILYIGTWLIDSEALHPLGMAVKSXSRGAGGARBEA 475  
 D 421 ADAGAGATAGIHMYSOILYIGTWLIDSEALHPLGMAVKSXSRGAGGARBEA 475

RESULT 6  
 AAY95977  
 ID AAY95977 standard; Protein: 475 AA.  
 AC AAY95977;  
 DT 05-DEC-2000 (first entry)  
 DE Human Sonic hedgehog Shh protein.  
 XX  
 KW Sonic hedgehog; Shh; human; agonist; antagonist;  
 KW lipid modulator; vacuole; cholesterol disorder; lipid disorder;  
 KW lipid metabolism; lipid storage; lipid transport; apolipoprotein;  
 KW triglyceride; hypercholesterolemia; abetalipoproteinemia;  
 KW hypobetalipoproteinemia; chylomicron retention; Anderson's disease;  
 KW fat absorption; atherosclerosis; obesity; weight loss;  
 KW vitamin A disorder; vitamin E disorder; anillopemia; anorectic;  
 KW antiarteriosclerotic; gene therapy; diagnosis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 463 /note="encoded by NNN"  
 XX  
 PN MO200051628-A2.  
 XX  
 PD 08-SEP-2000.  
 XX  
 PF 03-MAR-2000; 2000WO-US05662.  
 XX  
 PR 03-MAR-1999; 99US-0122640.  
 XX  
 PR 15-MAR-1999; 99US-0124446.  
 XX  
 PA (BIOU ) BIOGEN INC.

XX  
 PI Burkly L, Wang LC;  
 XX  
 DR WPI: 2000-611340/58.  
 DR N-PSDB; AAA95977.  
 XX  
 PT use of lipid modulators (e.g. hedgehog agonists or antagonists) for  
 PT modulating lipid metabolism and storage, especially useful for treating  
 PT lipid metabolism or cholesterol disorders, e.g. obesity or  
 PT hypercholesterolemia -  
 PS Disclosure; Page 118-120; 136pp; English.  
 XX  
 CC The present sequence of that of human Sonic hedgehog (Shh)  
 CC protein. The invention provides claimed methods for modulating  
 CC lipid metabolism, for modulating vacuole formation in intestinal  
 CC epithelial cells, for modulating the accumulation of fat in  
 CC intestinal epithelial cells, for treating a cholesterol disorder  
 CC and for treating a lipid metabolism disorder in an animal  
 CC (especially a human) by administering a lipid modulator selected  
 CC from a hedgehog antagonist or hedgehog agonist. In particular, the  
 CC lipid metabolism disorder is a lipid storage disorder, a lipid  
 CC transport disorder, an apolipoprotein disorder, a triglyceride  
 CC disorder, e.g. a triglyceride metabolism disorder, a triglyceride  
 CC transport disorder or a triglyceride storage disorder; a  
 CC diet-induced hypercholesterolemia, hypercholesterolemia,  
 CC abetalipoproteinemia, hypobetalipoproteinemia; a  
 CC chylomicron-retention disorder, Anderson's disease, a fat  
 CC absorption disorder, e.g. obesity or associated with weight loss,  
 CC normotriglyceridemic abetalipoproteinemia, an apolipoprotein-B100  
 CC deficiency, a fat soluble vitamin disorder, where the fat soluble  
 CC vitamin is vitamin A or E, or atherosclerosis (all claimed). The  
 CC hedgehog antagonist binds to the hedgehog receptor, but does not  
 CC elicit a response. It is preferably a hedgehog mimetic, a modified  
 CC hedgehog protein, e.g. an inactive hedgehog variant, or an  
 CC anti-hedgehog homologue, especially a human, chimeric or humanised  
 CC antibody. The methods are useful in preventing disorders or  
 CC protecting a subject from these disorders. The hedgehog antagonist  
 CC and agonist are also useful in diagnosis and research associated  
 CC with these disorders. The lipid modulators may also be used as a  
 CC part of a gene therapy protocol to deliver polynucleotides encoding  
 CC these lipid modulators.  
 CC  
 CC Sequence 475 AA;  
 XX  
 SQ

Query Match 99.9%; Score 2467; DB 21; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACPGRCGRRPKKLPPLAVAKOPIPNVAEKTGLASG 60  
 D 1 MLLARCLLVVSSLLVCSGLACPGRCGRRPKKLPPLAVAKOPIPNVAEKTGLASG 60  
 QY 61 RYEGKISNSERFKELPNPNPDIIFFDEENTGADRLMTORCKKLNALAISSVMNQPGV 120  
 D 61 RYEGKISNSERFKELPNPNPDIIFFDEENTGADRLMTORCKKLNALAISSVMNQPGV 120  
 QY 121 KLRVTEGDEDDHSEESLHYEGRAVDITTSDRRSKYGMARLAVEGPDWYYESKAH 180  
 D 121 KLRVTEGDEDDHSEESLHYEGRAVDITTSDRRSKYGMARLAVEGPDWYYESKAH 180  
 QY 181 IHCSVKAENSVAAKSGCGFPGSATVHLEOGCTKLVDLSPEDRYLAADDGRLLYSDFLT 240  
 D 181 IHCSVKAENSVAAKSGCGFPGSATVHLEOGCTKLVDLSPEDRYLAADDGRLLYSDFLT 240  
 QY 241 FLDDDDGAKKVFYIETREPERELLTPAHLFVAPHNDSATGPEASSSGSPSGALG 300  
 D 241 FLDDDDGAKKVFYIETREPERELLTPAHLFVAPHNDSATGPEASSSGSPSGALG 300  
 QY 301 PRALFASVRPGORVYVAERDGRRLTPAAVHSVTLSEEAAGAYAPLTAGTLLINRVL 360  
 D 301 PRALFASVRPGORVYVAERDGRRLTPAAVHSVTLSEEAAGAYAPLTAGTLLINRVL 360

OY 361 ASCYAVIEEHSMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
 DB 361 ASCYAVIEEHSMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
 OY 421 ADAPGAGATAGIHWYSOLLYOIGTWLIDSEALHPLGMAYKSSXSRRAGGAREGA 475  
 DB 421 ADAPGAGATAGIHWYSOLLYOIGTWLIDSEALHPLGMAYKSSXSRRAGGAREGA 475  
 RESULT 7  
 AAY95286  
 ID AAY95286 standard; Protein; 475 AA.  
 AC AAY95286;  
 XX 12-SEP-2000 (first entry)  
 DE Human Sonic hedgehog Shh protein.  
 KW Sonic hedgehog; Shh; human; excitotoxicity; Parkinson's disease;  
 KW Huntington's disease; neuronal degeneration; neuroprotective;  
 KW dopaminergic; GABAergic; substantia nigra; therapy.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 463 /note= "encoded by NNN"  
 FT  
 XX WO200035948-A1.  
 PD 22-JUN-2000.  
 PE 03-DEC-1999; 99WO-US28721.  
 PR 03-DEC-1998; 98WO-US25676.  
 PR 27-JAN-1999; 99US-0238243.  
 PR 03-JUN-1999; 99US-0325602.  
 PA (BIOJ ) BIOGEN INC.  
 PA (ONTO-) ONTOGENY INC.  
 PI Galdes A, Mahantappa N;  
 DR WPI: 2000-431570/37.  
 DR N-PSDB: AAA27881.  
 PT Treating disorders involving excitotoxicity, e.g. trauma, hypoglycemia,  
 PT senile dementia and Korsakoff's disease, by using lipophilic modified  
 PT hedgehog polypeptide -  
 PS Disclosure: Page 157-159; 174pp; English.  
 XX  
 CC The present sequence of that of human Sonic hedgehog (Shh)  
 CC protein. The invention relates to a method for promoting the  
 CC survival and/or functional performance of neuronal cells,  
 CC especially substantia nigra, dopaminergic or GABAergic neurons that  
 CC are susceptible to excitotoxicity, by contacting the cells, in vitro  
 CC or in vivo, with a lipophilic (e.g. cholesterol) modified hedgehog  
 CC polypeptide. The method is used to treat or prevent Parkinson's  
 CC disease, Huntington's disease, domoic acid poisoning, spinal cord  
 CC trauma, hypoglycemia, mechanical trauma to the nervous system,  
 CC senile dementia, Korsakoff's disease, schizophrenia, AIDS dementia,  
 CC multi-infarct dementia, mood disorders, depression, chemical  
 CC toxicity, neuronal damage associated with uncontrolled seizures  
 CC such as epileptic seizures, neuronal injury associated with HIV and  
 CC AIDS, neurodegeneration associated with Down's syndrome,  
 CC neuropathic pain syndrome, olivopontocerebral atrophy, amyotrophic  
 CC lateral sclerosis, mitochondrial abnormalities, Alzheimer's disease,  
 CC hepatic encephalopathy, Tourette's syndrome and drug addiction (all  
 CC claimed). The lipophilic modified hedgehog polypeptide is also  
 CC useful for promoting survival and/or functional performance of  
 CC neuronal cells susceptible to excitotoxicity.

XX SQ Sequence 475 AA;  
 Query Match 99.9%; Score 2467; DB 21; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 MLILARCLLVVSSLLVSGSLACGPGRCFGRRHPRKLTPLAYKQFIVNAEKTIGASG 60  
 DB 1 MLILARCLLVVSSLLVSGSLACGPGRCFGRRHPRKLTPLAYKQFIVNAEKTIGASG 60  
 OY 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKINALAISVMNQPCV 120  
 DB 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKINALAISVMNQPCV 120  
 OY 121 KLRVTEGMDGDHHSSESLHYEGRAVDITSDRDSKYGMLRLAVEAGFDVYYSKAH 180  
 DB 121 KLRVTEGMDGDHHSSESLHYEGRAVDITSDRDSKYGMLRLAVEAGFDVYYSKAH 180  
 OY 181 IHCYKAENSVAKSGCGPFSATVHLEGGTKLVKDLSPGDRVLAADQGRILYSDFLT 240  
 DB 181 IHCYKAENSVAKSGCGPFSATVHLEGGTKLVKDLSPGDRVLAADQGRILYSDFLT 240  
 OY 241 FLDRDDGAKKVVYIETREPRRLTLTAHLLFVAPHNDSATGEPBASSGSPSGALG 300  
 DB 241 FLDRDDGAKKVVYIETREPRRLTLTAHLLFVAPHNDSATGEPBASSGSPSGALG 300  
 OY 301 PRALFASRVPRQRYVYVVERDGRRLPLAAVHSVTLSEEAAGAVAPLTAQGTILINRYL 360  
 DB 301 PRALFASRVPRQRYVYVVERDGRRLPLAAVHSVTLSEEAAGAVAPLTAQGTILINRYL 360  
 OY 361 ASCYAVIEEHSMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
 DB 361 ASCYAVIEEHSMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGGRVALTAPGA 420  
 OY 421 ADAPGAGATAGIHWYSOLLYOIGTWLIDSEALHPLGMAYKSSXSRRAGGAREGA 475  
 DB 421 ADAPGAGATAGIHWYSOLLYOIGTWLIDSEALHPLGMAYKSSXSRRAGGAREGA 475  
 RESULT 8  
 AAY96248  
 ID AAY96248 standard; Protein; 475 AA.  
 AC AAY96248;  
 XX 11-SEP-2000 (first entry)  
 DE Partial human Shh.  
 KW Human; sonic hedgehog; Shh; neuromuscular disorder; neuropathy;  
 KW Guillain-Barre syndrome; peripheral neuropathy; diabetes; alcoholism;  
 KW chronic inflammatory demyelinating polyneuropathy; CIDP;  
 KW gene therapy; infection; inflammation; hereditary neuropathy;  
 KW Charcot-Marie-Tooth disease; vasculitis; lung cancer; tumour;  
 KW multiple myeloma; nutritional imbalance; kidney disease;  
 KW hypothyroid neuropathy; trauma; Refsum's disease; Abetalipoproteinemia;  
 KW Tanglefoot disease; Krabbe's disease; Metachromatic leukodystrophy;  
 KW Fabry's disease; CMT; GBS; Dejerine-Sottas syndrome; acute neuropathy;  
 KW amyotrophic lateral sclerosis; ALS; Miller-Fisher syndrome; amyloidosis;  
 KW hereditary sensory neuropathy type II; HSN II; B-cell lymphoma;  
 KW Waldenstrom's Macroglobulinemia; Chronic lymphocytic leukemia;  
 KW neuroprotective; cytoprotective; patched-mediated signal transduction.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 463 /label= unknown  
 FT /note= "Encoded by NNN"  
 FT  
 XX WO200027422-A2.





XX Sequence 475 AA: (99.9%) Score 2467; DB 21; Length 475;  
 Query Match Best Local Similarity 100.0%; Pred. No. 5.9e-218;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHPKLTPLAYKOFIPNVAEKTIGASG 60  
 DB 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHPKLTPLAYKOFIPNVAEKTIGASG 60  
 QY 61 RREGKISRNSERFKELTPYNPDIIFKDEENTGADRLMQRCKDKINALAISVMNQPGV 120  
 DB 61 RREGKISRNSERFKELTPYNPDIIFKDEENTGADRLMQRCKDKINALAISVMNQPGV 120  
 QY 121 KLRTVEGMDEGHSESLHYEGRAVDITTSRDRSKYGMRLAYEAGFDWVYTESKAH 180  
 DB 121 KLRTVEGMDEGHSESLHYEGRAVDITTSRDRSKYGMRLAYEAGFDWVYTESKAH 180  
 QY 181 IHCYKAENSVAKSGGCPGSATVHLEGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
 DB 181 IHCYKAENSVAKSGGCPGSATVHLEGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
 QY 241 FLDRODGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300  
 DB 241 FLDRODGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300  
 QY 301 PRALFASRYRPGORYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360  
 DB 301 PRALFASRYRPGORYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360  
 QY 361 ASCYAVIEHSHMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420  
 DB 361 ASCYAVIEHSHMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420  
 QY 421 AAPAGAGATAGIHWSQLLYOIGTWLDEALHPGLMAVKSXSXRGAGGARREGA 475  
 DB 421 AAPAGAGATAGIHWSQLLYOIGTWLDEALHPGLMAVKSXSXRGAGGARREGA 475

RESULT 10  
 AAG65748  
 ID AAG65748 standard; Protein; 475 AA.  
 AC AAG65748;  
 DT 07-JAN-2002 (first entry)  
 DE Human sonic hedgehog (Shh) polypeptide.  
 KW Adipocyte; hedgehog polypeptide; desert hedgehog; Indian hedgehog; Dhh;  
 KW Ihn; sonic hedgehog; Shh; therapeutic; cytostatic; tumor.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc-difference 463 /note="encoded by NNN"  
 XX WO200164238-A2.  
 PD 07-SEP-2001.  
 PF 28-FEB-2001; 2001WO-US06450.  
 PR 29-FEB-2000; 2000US-186058P.  
 PA (CURTI-) CURIS INC.  
 XX Zehentner B, Leser-Reliff U, Burtischer H;  
 XX WPI: 2001-607352/69.  
 DR N-PSDB; AAI66776.

XX Method for regulating formation and/or maintenance of adipocyte tissue  
 PT by contacting pre-adipocyte or adipocyte cells with a hedgehog  
 PI polypeptide or plc therapeutic  
 PS Disclosure; Page 102-104; 132pp; English.

XX The invention provides a method for regulating formation and/or  
 CC maintenance of adipocyte tissue that comprises contacting pre adipocyte  
 CC or adipocyte cells with a hedgehog polypeptide or plc therapeutic. The  
 CC method is used for regulating the growth state of an adipocyte stem/  
 CC progenitor cell, and treating or preventing disorders of, or surgical or  
 CC cosmetic repair of, adipocyte tissues, e.g. for treating or preventing  
 CC hyperplastic or neoplastic conditions affecting adipocyte tissue, such  
 CC as soft tissue tumors, especially adipose cell tumors, e.g. lipomas,  
 CC fibrolipomas, lipoblastomas, lipomatosis, hibernomas, hemangiomas and/or  
 CC liposarcomas. Hedgehog polypeptides can be used in combination with other  
 CC therapeutic agents. The present sequence represents a human sonic  
 CC hedgehog (Shh) polypeptide.

XX Sequence 475 AA: (99.9%) Score 2467; DB 22; Length 475;  
 Query Match Best Local Similarity 100.0%; Pred. No. 5.9e-218;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHPKLTPLAYKOFIPNVAEKTIGASG 60  
 DB 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHPKLTPLAYKOFIPNVAEKTIGASG 60  
 QY 61 RREGKISRNSERFKELTPYNPDIIFKDEENTGADRLMQRCKDKINALAISVMNQPGV 120  
 DB 61 RREGKISRNSERFKELTPYNPDIIFKDEENTGADRLMQRCKDKINALAISVMNQPGV 120  
 QY 121 KLRTVEGMDEGHSESLHYEGRAVDITTSRDRSKYGMRLAYEAGFDWVYTESKAH 180  
 DB 121 KLRTVEGMDEGHSESLHYEGRAVDITTSRDRSKYGMRLAYEAGFDWVYTESKAH 180  
 QY 181 IHCYKAENSVAKSGGCPGSATVHLEGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
 DB 181 IHCYKAENSVAKSGGCPGSATVHLEGGTKLVKDLSPGDRVLAADQGRLLYSDFLT 240  
 QY 241 FLDRODGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300  
 DB 241 FLDRODGAKKVYVYETREPRERLLTAHLLFVAPHNDSATGEPASSGSGPPSGGALG 300  
 QY 301 PRALFASRYRPGORYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360  
 DB 301 PRALFASRYRPGORYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360  
 QY 361 ASCYAVIEHSHMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420  
 DB 361 ASCYAVIEHSHMAHRAFAFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420  
 QY 421 AAPAGAGATAGIHWSQLLYOIGTWLDEALHPGLMAVKSXSXRGAGGARREGA 475  
 DB 421 AAPAGAGATAGIHWSQLLYOIGTWLDEALHPGLMAVKSXSXRGAGGARREGA 475

RESULT 11  
 AAB84674  
 ID AAB84674 standard; Protein; 475 AA.  
 AC AAB84674;  
 DT 17-SEP-2001 (first entry)  
 DE Amino acid sequence of a human hedgehog (Shh) polypeptide.  
 KW Shh gene; hedgehog gene; T lymphocyte; patched gene; infection;  
 KW diabetes; nutritional deficiency; graft rejection; hypercure response;  
 KW cornea transplant; autoimmune disorder; multiple sclerosis; psoriasis;  
 KW atopic dermatitis; inflammatory disease; proliferative disease;



KM hyperproliferative disease; eczematous dermatitis; urticaria;  
 KW vasculitis; scleroderma.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 463  
 TT /note- "unspecified residue encoded by NNN"  
 XX  
 PN MO200140438-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 30-NOV-2000; 2000MO-US32590.  
 XX  
 PR 30-NOV-1999; 99US-0168112.  
 XX  
 PA (CUR1-) CURIS INC.  
 XX  
 PI Crompton T;  
 XX  
 DR WPI; 2001-441484/47.  
 XX  
 DR N-PSDB; AAH28451.  
 XX  
 PT Modulating immune function comprises administration of a hedgehog or  
 PT ptc agent, for treating e.g. diabetes, eczematous dermatitis, urticaria  
 PT or vasculitis -  
 XX  
 PS Claim 4; Page 84-86; 105pp; English.  
 XX  
 CC The present sequence represents a hedgehog (Shh) polypeptide. Hedgehog  
 CC gene products and signal transduction pathways involving hedgehog are  
 CC involved in the maturation of T lymphocytes. The specification describes  
 CC a method for modulating immune function, by administration of a hedgehog  
 CC or patched (ptc) polypeptide, agonists or antagonists. The method is  
 CC used to treat disorders affecting the regulation of lymphocytes,  
 CC particularly maturation and/or activation of T lymphocytes, It is used,  
 CC to treat bacterial or viral infection, diabetes, nutritional  
 CC deficiencies, graft rejection or other hyperacute response such as  
 CC kidney, heart, lung, bone marrow spleen skin or cornea transplant or  
 CC autoimmune disorders such as multiple sclerosis, psoriasis or atopic  
 CC dermatitis. The method is used to treat inflammatory, proliferative and  
 CC hyperproliferative diseases, as well as cutaneous manifestations of  
 CC immunological disorders such as eczematous dermatitis, urticaria,  
 CC vasculitis and scleroderma.  
 CC  
 CC  
 SQ Sequence 475 AA;  
 Query Match 99.98; Score 2467; DB 22; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-218;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLLARCLLVIVSSLVCSGLAGRGRCRKRHRPKKLPPLAKKOFIPNVAEKTIGASG 60  
 DB 1 MLLARCLLVIVSSLVCSGLAGRGRCRKRHRPKKLPPLAKKOFIPNVAEKTIGASG 60  
 QY 61 RYEGKISRNSERFELTPNYPDIIFKDEENTGADRLMTORCKDKLALATISVNNOMPGV 120  
 DB 61 RYEGKISRNSERFELTPNYPDIIFKDEENTGADRLMTORCKDKLALATISVNNOMPGV 120  
 QY 121 KLRTYEGWDEDEGHSESLHTEGRAVDITTSDDRSKYKGMARLAVENAGDPWYTESKAH 180  
 DB 121 KLRTYEGWDEDEGHSESLHTEGRAVDITTSDDRSKYKGMARLAVENAGDPWYTESKAH 180  
 QY 181 IHCSVKAENSVAAKSGGCFPGSATVHLEOGGTLVKDLSPEDRVLAADDGRLLYSDFLT 240  
 DB 181 IHCSVKAENSVAAKSGGCFPGSATVHLEOGGTLVKDLSPEDRVLAADDGRLLYSDFLT 240  
 QY 241 FLDRDDCAKKVFYVETREPERRLTLAAHLFLVAPHNDSATGPEASSGSGPPSGGALG 300  
 DB 241 FLDRDDCAKKVFYVETREPERRLTLAAHLFLVAPHNDSATGPEASSGSGPPSGGALG 300  
 QY 301 PRALFAVRPGQRYVVAERDGRLLPAAVHSVTUSEAAGVAPLTAOGTILLINRYL 360

DB 301 PRALFAVRPGQRYVVAERDGRLLPAAVHSVTUSEAAGVAPLTAOGTILLINRYL 360  
 QY 361 ASCYAVTEESHMAHRAFPRLAALAPARTDGGSGGDRGGGRVALTPAGA 420  
 DB 361 ASCYAVTEESHMAHRAFPRLAALAPARTDGGSGGDRGGGRVALTPAGA 420  
 QY 421 ADAPGAGATAGIHWSQLYQIGTWLDSBALHPILGNAVSSXSRCGAGGAREGA 475  
 DB 421 ADAPGAGATAGIHWSQLYQIGTWLDSBALHPILGNAVSSXSRCGAGGAREGA 475  
 RESULT 12  
 AAED04687  
 ID AAED04687 standard; Protein; 475 AA.  
 XX  
 AC AAED04687;  
 XX  
 DT 04-SEP-2001 (first entry)  
 XX  
 DE Human sonic hedgehog (Shh) protein.  
 XX  
 KW Human; hedgehog protein; noctropic; neuroprotective; anticonvulsant;  
 KW cytoskeletal; therapy; Alzheimer's disease; Parkinson's disease; injury;  
 KW Huntington's chorea; amyotrophic lateral sclerosis; multiple sclerosis;  
 KW nervous system aging; neurodegenerative disease; immunological disease;  
 KW malignant glioma; medulloblastoma; neuroectodermal tumour; cancer;  
 KW extracellular signalling protein.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 24..197  
 FT /note- "N-terminal fragment"  
 FT Misc-difference 463  
 FT /label- Unknown  
 FT /note- "Encoded by NNN"  
 XX  
 PN MO200134654-A1.  
 XX  
 PD 17-MAY-2001.  
 XX  
 PF 02-NOV-2000; 2000MO-US30405.  
 XX  
 PR 05-NOV-1999; 99US-0164025.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI Strauch K;  
 XX  
 DR WPI; 2001-329075/34.  
 XX  
 DR N-PSDB; AAD09034.  
 XX  
 PT Novel isolated hedgehog fusion polypeptide useful for treating  
 PT neurological conditions such as Alzheimer's disease, Parkinson's  
 PT disease, Huntington's chorea, amyotrophic lateral sclerosis, and  
 PT multiple sclerosis -  
 XX  
 PS Disclosure; Page 117-118; 178pp; English.  
 XX  
 CC The present invention relates to hedgehog fusion proteins. Hedgehog  
 CC proteins are a family of extracellular signalling proteins that regulate  
 CC various aspects of embryonic development both in vertebrates and in  
 CC invertebrates. Hedgehog fusion protein is useful for the prophylaxis or  
 CC treatment of any condition or disease state for which a hedgehog or  
 CC patched protein constituent is efficacious and in the diagnosis of  
 CC constituents and/or conditions of disease states in non-physiological system or  
 CC specimens and for diagnostic purposes in non-physiological systems.  
 CC Hedgehog fusion protein is useful for treating neurological conditions  
 CC due to injury, aging of nervous system, including Alzheimer's disease,  
 CC chronic neurodegenerative diseases of the nervous system, including  
 CC Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis  
 CC and chronic immunological diseases of nervous system including multiple

CC sclerosts and malignant gliomas, medulloblastomas, neuroectodermal  
CC tumors and to specifically target medical therapies against cancers and  
CC tumors which express the receptor for the protein. The present sequence  
CC is human sonic hedgehog (Shh) protein.  
XX  
XX  
SQ Sequence 475 AA:

Query Match 99.9%; Score 2467; DB 22; Length 475;  
Best Local Similarity 100.0%; Pred. No. 5.9e-218;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MILLARCLLVVSSLLVCSGLACGPGFGRRHPRKLTPLAYKQFIPNVAEKTIGASG 60  
DB 1 MILLARCLLVVSSLLVCSGLACGPGFGRRHPRKLTPLAYKQFIPNVAEKTIGASG 60  
QY 61 RYEGKISRNSERFEKLTPTNPNPDIFKDEENTGADRLMTORCKDKLNALATISVMNQPGV 120  
DB 61 RYEGKISRNSERFEKLTPTNPNPDIFKDEENTGADRLMTORCKDKLNALATISVMNQPGV 120  
QY 121 KLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYESKAH 180  
DB 121 KLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYESKAH 180  
QY 181 HICSVKAENSVAAKSGCGPGSATVHLBOGCKTKLVKDLSPGDRVLAADQGLTSDFLT 240  
DB 181 HICSVKAENSVAAKSGCGPGSATVHLBOGCKTKLVKDLSPGDRVLAADQGLTSDFLT 240  
QY 241 FLDRDGAKKVYVETREPRERLLTAAHLTFVAPHNDSATGEPASSGSGPPSGALG 300  
DB 241 FLDRDGAKKVYVETREPRERLLTAAHLTFVAPHNDSATGEPASSGSGPPSGALG 300  
QY 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSTLSEEAAGAYAPLTAOGTILINRVL 360  
DB 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSTLSEEAAGAYAPLTAOGTILINRVL 360  
QY 361 ASCYAVIEHSNAHRAFAFRLAHLALAPARDRGDSCGPDGGGGVATLAPAA 420  
DB 361 ASCYAVIEHSNAHRAFAFRLAHLALAPARDRGDSCGPDGGGGVATLAPAA 420  
QY 421 ADAPGAGATAGIHWSQLLYQITWLLDSEALHPLOMAVKSXSXRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWSQLLYQITWLLDSEALHPLOMAVKSXSXRGAGGAREGA 475

RESULT 13  
AAE05377  
ID AAE05377 standard; Protein; 475 AA.

XX AAE05377;  
XX 12-SEP-2001 (first entry)  
DE Human Sonic hedgehog protein.  
XX  
XX Human; Sonic hedgehog; Shh; morphogenic signal; neuron; chromosome 7q;  
KW embryonic patterning; cell culture; cell differentiation; ischaemia;  
KW cell proliferative disorder; intracerebral grafting; Huntington's chorea;  
KW neurological disorder; Alzheimer's disease; Parkinson's disease;  
KW amyotrophic lateral sclerosis; ALS; multiple sclerosis.

XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Misc-difference 463 /label= Unknown  
XX FT /note= "Encoded by NNN"  
XX  
XX US6261786-B1.  
XX 17-JUL-2001.  
XX 02-JUL-1996; 96US-0674509.

PR 30-DEC-1993; 93US-0176427.  
PR 14-DEC-1994; 94US-0356060.  
PR 04-MAY-1995; 95US-0435093.  
PR 05-JUN-1995; 95US-0460900.  
PR 05-JUN-1995; 95US-0462386.

XX  
XX (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
XX (HARD ) HARVARD COLLEGE.

PA Marigo V, Tablin CJ, Ingham PW, McMahon AP;

DR WPI. 2001-440859/47.  
DR N-PSDB: AAD10151.

PT Screening compounds that potentiate or inhibit binding of hedgehog  
PT polypeptide to naturally occurring patched receptor, comprises  
PT contacting polypeptide with receptor and test compound, and detecting  
PT change in binding

PS Claim 2; Column 163-166; 127pp; English.

CC The present invention relates to assay for screening compounds that  
CC potentiate or inhibit binding of hedgehog polypeptide to naturally  
CC occurring patched receptor. The hedgehog proteins comprise morphogenic  
CC signals produced by embryonic patterning centres, and are involved in the  
CC formation and maintenance of ordered spatial arrangements of  
CC differentiated tissues in vertebrates, both adult and embryonic. The  
CC proteins can be used to generate and/or maintain an array of different  
CC vertebrate tissues both in vitro and in vivo. The invention also relates  
CC to a method for modulating growth, differentiation or survival of a  
CC mammalian cell (e.g. neuron, testicular cell) responsive to hedgehog  
CC induction. Hedgehog agonists and antagonists can be used in cell culture  
CC techniques to enhance survival and maintenance of neurons and various  
CC vertebrate organogenic pathways. The hedgehog gene is useful in  
CC determining whether a patient is at the risk of disorder characterised by  
CC unwanted cell proliferation or aberrant control of differentiation. The  
CC hedgehog proteins or mimetics can be used to induce foetal neurons  
CC especially neuronal stem cells in intracerebral grafting. The protein  
CC or its mimetic can be used in the treatment of neurological conditions  
CC e.g. injury to nervous system, ischaemia resulting from stroke,  
CC Alzheimer's disease, Parkinson's disease, Huntington's chorea,  
CC amyotrophic lateral sclerosis (ALS) and multiple sclerosis. The present  
CC sequence is human sonic hedgehog (Shh) protein. The Shh gene is  
CC located on human chromosome 7q.  
XX

SQ Sequence 475 AA:

Query Match 99.9%; Score 2467; DB 22; Length 475;  
Best Local Similarity 100.0%; Pred. No. 5.9e-218;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MILLARCLLVVSSLLVCSGLACGPGFGRRHPRKLTPLAYKQFIPNVAEKTIGASG 60  
DB 1 MILLARCLLVVSSLLVCSGLACGPGFGRRHPRKLTPLAYKQFIPNVAEKTIGASG 60  
QY 61 RYEGKISRNSERFEKLTPTNPNPDIFKDEENTGADRLMTORCKDKLNALATISVMNQPGV 120  
DB 61 RYEGKISRNSERFEKLTPTNPNPDIFKDEENTGADRLMTORCKDKLNALATISVMNQPGV 120  
QY 121 KLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYESKAH 180  
DB 121 KLRVTEGMDGDHSEESLHYEGRAVDITTSRDRSKYGMRLAVEAGFDWVYESKAH 180  
QY 181 HICSVKAENSVAAKSGCGPGSATVHLBOGCKTKLVKDLSPGDRVLAADQGLTSDFLT 240  
DB 181 HICSVKAENSVAAKSGCGPGSATVHLBOGCKTKLVKDLSPGDRVLAADQGLTSDFLT 240  
QY 241 FLDRDGAKKVYVETREPRERLLTAAHLTFVAPHNDSATGEPASSGSGPPSGALG 300  
DB 241 FLDRDGAKKVYVETREPRERLLTAAHLTFVAPHNDSATGEPASSGSGPPSGALG 300  
QY 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSTLSEEAAGAYAPLTAOGTILINRVL 360

Db 301 PRLAFASRVPRGQRYVVAERDGRLLPAAVHSVTLSEBAAGAYAPLTAGTILLINRVL 360  
 QY 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGDSGGDRGGCGGVALTPAGA 420  
 Db 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGDSGGDRGGCGGVALTPAGA 420  
 QY 421 ADAPGAGATGIMHYSOLLQIGTWLDSALHPLGMAVSSXSRCAGGARBCGA 475  
 Db 421 ADAPGAGATGIMHYSOLLQIGTWLDSALHPLGMAVSSXSRCAGGARBCGA 475  
 RESULT 14  
 ID AAB31222 standard; Protein: 475 AA.  
 XX AAB31222;  
 XX 20-APR-2001 (first entry)  
 DE Amino acid sequence of human sonic hedgehog protein (Shh).  
 XX Hedgehog related-protein; sonic hedgehog protein; Shh; Ischemia; stroke;  
 KM desert hedgehog protein; Dhh; Indian hedgehog protein; Ihh; neuron;  
 KM neurological condition; nervous system injury; tumour-induced injury;  
 KM aging; Alzheimer's disease; chronic neurodegenerative disease;  
 KM Parkinson's disease; Huntington's chorea; amyotrophic lateral sclerosis;  
 KM spinocerebellar degeneration; chronic immunological disease;  
 KM multiple sclerosis.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT MISC-difference 463  
 FT /note="unspecified amino acid encoded by NNN"  
 XX US6155747-A.  
 XX 26-DEC-2000.  
 PF 05-JUN-1995; 95US-0460900.  
 XX 30-DEC-1993; 93US-0176427.  
 PR 14-DEC-1994; 94US-0356060.  
 PR 04-MAY-1995; 95US-0435093.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.  
 PI Ingham PW, McMahon AB, Tablin CJ, Marti-gorostiza E, Bumcrot DA;  
 XX N-PSDB; AAC87079.  
 DR WPI: 2001-079847/09.  
 XX  
 PT Polynucleotides encoding hedgehog proteins, useful for treating  
 PT diseases of nervous system such as Alzheimer's disease, Parkinson's  
 PT disease, Huntington's chorea, amyotrophic lateral sclerosis, multiple  
 PT sclerosis -  
 PS Claim 10: Columns 147-150; 11pp; English.  
 XX The present sequence represents a hedgehog related-protein. The  
 CC specification describes a sonic hedgehog protein (Shh), a desert  
 CC hedgehog protein (Dhh), and an Indian hedgehog protein (Ihh). The  
 CC hedgehog polynucleotides are useful in diagnostic, in antisense  
 CC therapy and in therapeutic assays for detecting and treating disorders  
 CC involving, e.g., aberrant expression of vertebrate hedgehog homologues.  
 CC hedgehog polypeptides are useful therapeutically to enhance survival  
 CC of neurons and other neuron cells and in treating neurological  
 CC conditions deriving from acute, subacute, or chronic injury to the  
 CC nervous system, including traumatic injury, chemical injury, vascular  
 CC injury and deficits (such as the ischemia resulting from stroke),  
 CC together with infectious/inflammatory and induced-induced injury, aging  
 CC of the nervous system including Alzheimer's disease, chronic

CC neurodegenerative diseases of the nervous system, including Parkinson's  
 CC disease, Huntington's chorea, amyotrophic lateral sclerosis,  
 CC spinocerebellar degenerations, and chronic immunological diseases of  
 CC the nervous system or affecting the nervous system, including multiple  
 CC sclerosis.  
 XX  
 SQ Sequence 475 AA:  
 Query Match 99.9%; Score 2467; DB 22; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 5; 9e-218;  
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLLIARCLLLVLYSSLVCSGLACGPGRGFKRRHPKRLTPLAYKQFIPVAEKTIGASG 60  
 Db 1 MLLIARCLLLVLYSSLVCSGLACGPGRGFKRRHPKRLTPLAYKQFIPVAEKTIGASG 60  
 QY 61 RYEGKISRNSERREKELTPNNPDIIFFDEENTGADRLMTQCKDKNALISVNMOPGV 120  
 Db 61 RYEGKISRNSERREKELTPNNPDIIFFDEENTGADRLMTQCKDKNALISVNMOPGV 120  
 QY 121 KLRTYEGMDGDGHSESLHYEGRAVDITTSDBDRSKYGMILARLAVAGPDWYYESKAH 180  
 Db 121 KLRTYEGMDGDGHSESLHYEGRAVDITTSDBDRSKYGMILARLAVAGPDWYYESKAH 180  
 QY 181 IHCYKAENSVAAKSGGCFPGSATVHLEGGCTKLVKDLSFGDRYLADDDGRLLYSDFLT 240  
 Db 181 IHCYKAENSVAAKSGGCFPGSATVHLEGGCTKLVKDLSFGDRYLADDDGRLLYSDFLT 240  
 QY 241 FLDRDCAKKVFYIERERERERLLTAAHLLVPAHPNDATGEPKSSSGSPSGALG 300  
 Db 241 FLDRDCAKKVFYIERERERERLLTAAHLLVPAHPNDATGEPKSSSGSPSGALG 300  
 QY 301 PRLAFASRVPRGQRYVVAERDGRLLPAAVHSVTLSEBAAGAYAPLTAGTILLINRVL 360  
 Db 301 PRLAFASRVPRGQRYVVAERDGRLLPAAVHSVTLSEBAAGAYAPLTAGTILLINRVL 360  
 QY 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGDSGGDRGGCGGVALTPAGA 420  
 Db 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGDSGGDRGGCGGVALTPAGA 420  
 QY 421 ADAPGAGATGIMHYSOLLQIGTWLDSALHPLGMAVSSXSRCAGGARBCGA 475  
 Db 421 ADAPGAGATGIMHYSOLLQIGTWLDSALHPLGMAVSSXSRCAGGARBCGA 475  
 RESULT 15  
 ID AAB60265 standard; Protein: 475 AA.  
 XX AAB60265;  
 AC 30-MAR-2001 (first entry)  
 XX  
 DE Human Sonic hedgehog (Shh) protein, SEQ ID NO:15.  
 XX Hedgehog protein; polymer conjugate; polyalkene glycol group;  
 KM bioavailability; formulation; neurological disorder;  
 KM inflammatory disorder; autoimmune disorder; cancer;  
 KM neurodegenerative disorder; Parkinson's disease; Huntington's disease;  
 KM Alzheimer's disease; neurological injury; stroke; multiple sclerosis;  
 KM malignant glioma; medulloblastoma; neuroectodermal tumour.  
 OS Homo sapiens.  
 XX  
 PN WO200073337-A1.  
 XX  
 PD 07-DEC-2000.  
 XX  
 PF 26-MAY-2000; 2000MO-US14741.  
 XX  
 PR 01-JUN-1999; 99US-0137011.  
 PR 13-AUG-1999; 99US-0149016.  
 XX

PA (BIOT) BIOGEN INC.

XX Pepinsky RB, Taylor F, Garber E;

XX WPI: 2001-049927/06.

DR N-PSDB: AAF27018.

XX Modified hedgehog protein, useful in the treatment of Parkinson's  
PT disease and Huntington's chorea, comprises a polymer containing a  
PT polyalkylene glycol group linked to any residue other than the  
PT N-terminal and lysine residues -

XX Disclosure: Page 138-140; 157pp; English.

XX The invention relates to novel polymer conjugates of hedgehog proteins  
CC which have increased bioavailability. The hedgehog proteins are  
CC conjugated to a non-naturally-occurring polymer comprising a polyalkylene  
CC glycol group, with the proviso that the polymer is not conjugated to the  
CC N-terminus, or to lysine residues of the hedgehog protein. The hedgehog  
CC protein used in the conjugate may be a wild-type or mutant Sonic hedgehog  
CC (Shh), Indian hedgehog (Ihh) or Desert hedgehog (Dhh) protein, or may be  
CC a hedgehog fusion protein. The invention also relates to methods of  
CC defining and mapping functionally important regions of a protein by  
CC modifying accessible amino acid side chains, and determining the effect  
CC the position and/or type of modification have on the activity of the  
CC protein. The hedgehog polymer conjugates may be used in the management of  
CC various medical conditions including various neurological disorders,  
CC inflammatory and autoimmune diseases, and cancers. In particular, they  
CC may be used to prevent preventing or ameliorate neurodegenerative  
CC disorders (e.g., Parkinson's disease, Huntington's disease, Alzheimer's  
CC disease); age-associated neurological disease; neurological injury and  
CC trauma; immunological diseases of the nervous system (e.g., multiple  
CC sclerosis); stroke; and malignant gliomas, medulloblastomas and  
CC neuroectodermal tumors. The modifications made to the hedgehog protein  
CC may result in increased half-life, altered tissue distribution (such as  
CC an improved ability to stay in the vasculature for longer periods of  
CC time), increased stability in solution, protection from proteolytic  
CC degradation, or reduced immunogenicity. In particular, the ability to  
CC remain in the vasculature for prolonged periods may allow a hedgehog  
CC protein of the invention to cross the blood-brain barrier, and an  
CC increased thermal stability would be an advantage when formulating the  
CC hedgehog protein in powder form. The present sequence represents a  
CC member of the hedgehog family of proteins.

XX Sequence 475 AA:

Query Match 99.9%; Score 2467; DB 22; Length 475;

Best Local Similarity 100.0%; Pred. No. 5.9e-218;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHRRPKLTPLAYKQFIIPNVAEKTLAGSG 60  
DB 1 MLLARCLLVVSSLLVCSGLACGPGRGFRHRRPKLTPLAYKQFIIPNVAEKTLAGSG 60  
OY 61 RYEGKISRSEFEKELTPYNNPDIIFKDENTGADRLMTQRCCKLNLALISVNMQPGV 120  
DB 61 RYEGKISRSEFEKELTPYNNPDIIFKDENTGADRLMTQRCCKLNLALISVNMQPGV 120  
OY 121 KLRVTEGMDHSEESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDMVYVESKAH 180  
DB 121 KLRVTEGMDHSEESLHYEGRAVDITTSRDSKYGMLARLAVEAGFDMVYVESKAH 180  
OY 181 IHCYKAVNSVAAKSGGCPGSAVHLBEGGTLYKVLSPGDRVLAADQGRILYSDFLT 240  
DB 181 IHCYKAVNSVAAKSGGCPGSAVHLBEGGTLYKVLSPGDRVLAADQGRILYSDFLT 240  
OY 241 FLDRDDGAKKVFYVETREPRERLLTAHLFVAPPHNDSATGEPPASSGSPSGALG 300  
DB 241 FLDRDDGAKKVFYVETREPRERLLTAHLFVAPPHNDSATGEPPASSGSPSGALG 300  
OY 301 PRLAFASRYRPGORYVVAERDGRRLPAAVHSTLSEEAAGAYAPLTAQSTILLNRYL 360  
DB 301 PRLAFASRYRPGORYVVAERDGRRLPAAVHSTLSEEAAGAYAPLTAQSTILLNRYL 360

OY 361 ASCYAVIEEHSNAHRAFAFPRLAHALLALAPARTDRGDSGGDRGGGGRVALTPAGA 420

DB 361 ASCYAVIEEHSNAHRAFAFPRLAHALLALAPARTDRGDSGGDRGGGGRVALTPAGA 420

OY 421 ADAPGAGATAGIHWYSQLLYIGTWLIDSEALHPGLMAVKSXSRSRAGGAREGA 475

DB 421 ADAPGAGATAGIHWYSQLLYIGTWLIDSEALHPGLMAVKSXSRSRAGGAREGA 475

Search completed: February 20, 2003, 10:10:20  
Job time : 36.4615 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 20, 2003, 10:08:38 ; Search time 12.4346 Seconds  
(without alignments)  
1123.956 Million cell updates/sec

Title: US-09-827-110A-15  
Perfect score: 2469  
Sequence: 1 MLLARCLLVVSLVCS.....GNAVKSSRGAGGAREGA 475

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata1/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata1/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2467	99.9	475	2	US-08-356-060A-13
2	2467	99.9	475	4	US-08-460-900C-13
3	2467	99.9	475	4	US-08-674-509B-13
4	2467	99.9	475	4	US-08-954-698-13
5	2467	99.9	475	4	US-08-954-698-13
6	2467	99.9	475	4	US-08-957-874-13
7	2407	97.5	462	1	US-09-325-256-22
8	2407	97.5	462	1	US-08-748-591-4
9	2068	83.8	437	3	US-08-946-329A-20
10	2068	83.8	437	4	US-08-567-357A-20
11	2068	83.8	437	4	US-08-729-743A-20
12	2068	83.8	437	4	US-09-057-860A-6
13	2068	83.8	437	5	US-08-349-198-20
14	2068	83.8	437	5	PCT-US95-15463-19
15	2068	83.8	437	5	PCT-US95-15923-20
16	2065	83.6	437	1	US-08-176-427B-8
17	2065	83.6	437	2	US-08-356-060A-11
18	2065	83.6	437	4	US-08-460-900C-11
19	2065	83.6	437	4	US-08-674-509B-11
20	2065	83.6	437	4	US-08-954-698-11
21	2065	83.6	437	4	US-08-957-874-11
22	2065	83.6	437	4	US-09-325-256-20
23	2062	83.5	437	4	US-09-293-305-14
24	2060	83.4	437	4	US-08-757-230A-2
25	2060	83.4	437	4	US-08-757-230A-9
26	2060	83.4	437	5	PCT-US95-02315-2
27	1807	73.2	425	1	US-08-176-427B-2

28	1807	73.2	425	2	US-08-356-060A-8	Sequence 8, Appl1
29	1807	73.2	425	4	US-08-460-900C-8	Sequence 8, Appl1
30	1807	73.2	425	4	US-08-674-509B-8	Sequence 8, Appl1
31	1807	73.2	425	4	US-08-954-698-8	Sequence 8, Appl1
32	1807	73.2	425	4	US-08-957-874-8	Sequence 8, Appl1
33	1807	73.2	425	4	US-09-325-256-17	Sequence 17, Appl1
34	1690	68.4	425	3	US-08-946-329A-19	Sequence 19, Appl1
35	1690	68.4	425	4	US-08-567-357A-19	Sequence 19, Appl1
36	1690	68.4	425	4	US-08-729-743A-19	Sequence 19, Appl1
37	1690	68.4	425	4	US-08-349-198-19	Sequence 19, Appl1
38	1690	68.4	425	5	PCT-US95-15463-19	Sequence 19, Appl1
39	1690	68.4	425	5	PCT-US95-15923-19	Sequence 19, Appl1
40	1506	61.0	418	3	US-08-946-329A-18	Sequence 18, Appl1
41	1506	61.0	418	4	US-08-567-357A-18	Sequence 18, Appl1
42	1506	61.0	418	4	US-08-729-743A-18	Sequence 18, Appl1
43	1506	61.0	418	4	US-08-757-230A-7	Sequence 7, Appl1
44	1506	61.0	418	4	US-08-349-198-18	Sequence 18, Appl1
45	1506	61.0	418	5	PCT-US95-15463-18	Sequence 18, Appl1

## ALIGNMENTS

RESULT 1  
US-08-356-060A-13  
Sequence 13, Application US/08356060A  
Patent No. 5844079

## GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
TITLE OF INVENTION: Tablin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/356,060A  
FILING DATE: 14-DEC-1994  
CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMI-006CP  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-356-060A-13

Query Match 99.9%; Score 2467; DB 2; Length 475;  
Best Local Similarity 100.0%; Pred. No. 8, 2e-239;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLLARCLLVVSLVCSGLACGPGRGKRRHKKLTPLAYKQFIPNVAEKTLGASG 60

|||||  
Db 1 MILLARCLLVYSSILVSGIACGPGFGRRHPKKTLPYAKOFTPNVAEKTGASG 60  
QY 61 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLMLAISVMNQPGV 120  
Db 61 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLMLAISVMNQPGV 120  
QY 121 KLRVTEGMDECHHSEESIHYEGRAVDITTSRDRSKYGMLARLAVEAGFDMVYESKAH 180  
Db 121 KLRVTEGMDECHHSEESIHYEGRAVDITTSRDRSKYGMLARLAVEAGFDMVYESKAH 180  
QY 181 HCSYKAENSVAKSGGCGPGSATVHLEGGTKLVKDLSPGDRVLAADOGRLTSDFLT 240  
Db 181 HCSYKAENSVAKSGGCGPGSATVHLEGGTKLVKDLSPGDRVLAADOGRLTSDFLT 240  
QY 241 FLDRDGAKKVYVYIETREPRERLLTAAHLTFVAPHNDATGEPEASSGSGPPSGALG 300  
Db 241 FLDRDGAKKVYVYIETREPRERLLTAAHLTFVAPHNDATGEPEASSGSGPPSGALG 300  
QY 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSTVLTSEAAAGATAPLTAGTIIINRYL 360  
Db 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSTVLTSEAAAGATAPLTAGTIIINRYL 360  
QY 361 ASCYAVIEESHVAHAFAPFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420  
Db 361 ASCYAVIEESHVAHAFAPFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420  
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLDESEALHPGMAVKSXSXSGAGGAREGA 475  
Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLDESEALHPGMAVKSXSXSGAGGAREGA 475

RESULT 2  
US-08-460-900C-13  
; Sequence 13, Application US/08460900C  
; Patent No. 6165747  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tablin, Clifford J.  
; APPLICANT: Bumcrot, David A.  
; APPLICANT: Marli-Gorositz, Elisa  
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
; TITLE OF INVENTION: Proteins and Uses Related Thereto  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,900C  
; FILING DATE: 5-JUNE-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,093  
; FILING DATE: 4-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/356,060  
; FILING DATE: 14-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/176,427  
; FILING DATE: 30-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: HWY-006.05

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 475 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-460-900C-13

Query Match 99.9%; Score 2467; DB 4; Length 475;  
Best Local Similarity 100.0%; Pred. No. 8.2e-239;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|||||  
QY 1 MILLARCLLVYSSILVSGIACGPGFGRRHPKKTLPYAKOFTPNVAEKTGASG 60  
Db 1 MILLARCLLVYSSILVSGIACGPGFGRRHPKKTLPYAKOFTPNVAEKTGASG 60  
QY 61 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLMLAISVMNQPGV 120  
Db 61 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLMLAISVMNQPGV 120  
QY 121 KLRVTEGMDECHHSEESIHYEGRAVDITTSRDRSKYGMLARLAVEAGFDMVYESKAH 180  
Db 121 KLRVTEGMDECHHSEESIHYEGRAVDITTSRDRSKYGMLARLAVEAGFDMVYESKAH 180  
QY 181 HCSYKAENSVAKSGGCGPGSATVHLEGGTKLVKDLSPGDRVLAADOGRLTSDFLT 240  
Db 181 HCSYKAENSVAKSGGCGPGSATVHLEGGTKLVKDLSPGDRVLAADOGRLTSDFLT 240  
QY 241 FLDRDGAKKVYVYIETREPRERLLTAAHLTFVAPHNDATGEPEASSGSGPPSGALG 300  
Db 241 FLDRDGAKKVYVYIETREPRERLLTAAHLTFVAPHNDATGEPEASSGSGPPSGALG 300  
QY 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSTVLTSEAAAGATAPLTAGTIIINRYL 360  
Db 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSTVLTSEAAAGATAPLTAGTIIINRYL 360  
QY 361 ASCYAVIEESHVAHAFAPFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420  
Db 361 ASCYAVIEESHVAHAFAPFRLAHALLAALAPARTDRGDSGGGGRGVALTAPGA 420  
QY 421 ADAPGAGATAGIHWYSQLLYQIGTWLDESEALHPGMAVKSXSXSGAGGAREGA 475  
Db 421 ADAPGAGATAGIHWYSQLLYQIGTWLDESEALHPGMAVKSXSXSGAGGAREGA 475

RESULT 3  
US-08-674-509B-13  
; Sequence 13, Application US/08674509B  
; Patent No. 6261786  
; GENERAL INFORMATION:  
; APPLICANT: Ingham, Phillip W.  
; APPLICANT: McMahon, Andrew P.  
; APPLICANT: Tablin, Clifford J.  
; APPLICANT: Marigo, Valeria  
; TITLE OF INVENTION: SCREENING ASSAYS FOR HEDGCG AGONISTS  
; TITLE OF INVENTION: AND ANTAGONISTS  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/674,509B  
FILING DATE: 02-JUL-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/460,900  
FILING DATE: 05-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006,06  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-674-509B-13

Query Match 99.9% Score 2467; DB 4; Length 475;  
Best Local Similarity 100.0%; Pred. No. 8.2e-239;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRRHPRKPLPLAYKQFIPIVAEKTIGASG 60  
DB 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRRHPRKPLPLAYKQFIPIVAEKTIGASG 60  
QY 61 RYEGKISNSERFELPNYNDIIFKDEENTGADRLMTORCKDLNALAISVNMQMPGV 120  
DB 61 RYEGKISNSERFELPNYNDIIFKDEENTGADRLMTORCKDLNALAISVNMQMPGV 120  
QY 121 KLRTGEGDEGHHSESLHTEGRAVDITTSDRRSKYGMRLAVEGFMVYYESKAH 180  
DB 121 KLRTGEGDEGHHSESLHTEGRAVDITTSDRRSKYGMRLAVEGFMVYYESKAH 180  
QY 181 IHCSVKAENSVAAKSGCGFPGSATVHLEQGCTKLVKDLSPGDRVLAADDGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAAKSGCGFPGSATVHLEQGCTKLVKDLSPGDRVLAADDGRLLYSDFLT 240  
QY 241 FLRDDGAKKVFYIETREPERLLTAAHLLEFVAAPHNDSATGEPASSSGSPSGALG 300  
DB 241 FLRDDGAKKVFYIETREPERLLTAAHLLEFVAAPHNDSATGEPASSSGSPSGALG 300  
QY 301 PRALEASRVRRGORYVVAERDGRRLPAVHSTLSEEAAGVAPLTAQGTILINRVL 360  
DB 301 PRALEASRVRRGORYVVAERDGRRLPAVHSTLSEEAAGVAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEHSMARAFAPFLAHALLAALAPARTDRGDSGGDRGGGGRVALTPAGA 420  
DB 361 ASCYAVIEHSMARAFAPFLAHALLAALAPARTDRGDSGGDRGGGGRVALTPAGA 420  
QY 421 ADAPGAGTAGIHWYSOLLIOIGTWLDSEALHPLGMAVKSXSXRGAGGAREGA 475  
DB 421 ADAPGAGTAGIHWYSOLLIOIGTWLDSEALHPLGMAVKSXSXRGAGGAREGA 475

## RESULT 4

US-08-954-698-13  
Sequence 13, Application US/08954698

Patent No. 6271363  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tablin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
TITLE OF INVENTION: Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELLIOT LLP  
STREET: One Post Office Square  
CITY: Boston

STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/954,698  
FILING DATE: 20-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/462,386  
FILING DATE: 05-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 04-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006,10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-954-698-13

Query Match 99.9% Score 2467; DB 4; Length 475;  
Best Local Similarity 100.0%; Pred. No. 8.2e-239;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRRHPRKPLPLAYKQFIPIVAEKTIGASG 60  
DB 1 MLLARCLLVVSSLLVCSGLACGPGRGFKRRHPRKPLPLAYKQFIPIVAEKTIGASG 60  
QY 61 RYEGKISNSERFELPNYNDIIFKDEENTGADRLMTORCKDLNALAISVNMQMPGV 120  
DB 61 RYEGKISNSERFELPNYNDIIFKDEENTGADRLMTORCKDLNALAISVNMQMPGV 120  
QY 121 KLRTGEGDEGHHSESLHTEGRAVDITTSDRRSKYGMRLAVEGFMVYYESKAH 180  
DB 121 KLRTGEGDEGHHSESLHTEGRAVDITTSDRRSKYGMRLAVEGFMVYYESKAH 180  
QY 181 IHCSVKAENSVAAKSGCGFPGSATVHLEQGCTKLVKDLSPGDRVLAADDGRLLYSDFLT 240  
DB 181 IHCSVKAENSVAAKSGCGFPGSATVHLEQGCTKLVKDLSPGDRVLAADDGRLLYSDFLT 240  
QY 241 FLRDDGAKKVFYIETREPERLLTAAHLLEFVAAPHNDSATGEPASSSGSPSGALG 300  
DB 241 FLRDDGAKKVFYIETREPERLLTAAHLLEFVAAPHNDSATGEPASSSGSPSGALG 300  
QY 301 PRALEASRVRRGORYVVAERDGRRLPAVHSTLSEEAAGVAPLTAQGTILINRVL 360  
DB 301 PRALEASRVRRGORYVVAERDGRRLPAVHSTLSEEAAGVAPLTAQGTILINRVL 360  
QY 361 ASCYAVIEHSMARAFAPFLAHALLAALAPARTDRGDSGGDRGGGGRVALTPAGA 420  
DB 361 ASCYAVIEHSMARAFAPFLAHALLAALAPARTDRGDSGGDRGGGGRVALTPAGA 420  
QY 421 ADAPGAGTAGIHWYSOLLIOIGTWLDSEALHPLGMAVKSXSXRGAGGAREGA 475  
DB 421 ADAPGAGTAGIHWYSOLLIOIGTWLDSEALHPLGMAVKSXSXRGAGGAREGA 475

# RESULT 5 US-08-957-874-13

Sequence 13, Application US/08957874  
Patent No. 6384192

## GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tablin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
TITLE OF INVENTION: Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 47  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/957,874  
FILING DATE: 20-OCT-1997

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/462,386  
FILING DATE: 5-JUNE-1995

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/435,093  
FILING DATE: 4-MAY-1995

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993

## ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006,09

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000

## INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-957-874-13

Query Match 99.9%; Score 2467; DB 4; Length 475;  
Best Local Similarity 100.0%; Pred. No. 8.2e-239;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHKITPLAYKQFIPNVAEKTIGASG 60  
DB 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHKITPLAYKQFIPNVAEKTIGASG 60  
QY 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNQMPGV 120  
DB 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNQMPGV 120  
QY 121 KLRVTEGMDEDDHSEESLHYEGRAVDITTSRDRSKYGMGLARLAVEAGFDVYVESKAH 180  
DB 121 KLRVTEGMDEDDHSEESLHYEGRAVDITTSRDRSKYGMGLARLAVEAGFDVYVESKAH 180  
QY 181 IHCYKAENSVAAKSGGCGPGSATVHLEGGTKLVKDLSPGDRVLAADOGRLLYSDFLT 240  
DB 181 IHCYKAENSVAAKSGGCGPGSATVHLEGGTKLVKDLSPGDRVLAADOGRLLYSDFLT 240

DB 181 IHCYKAENSVAAKSGGCGPGSATVHLEGGTKLVKDLSPGDRVLAADOGRLLYSDFLT 240  
QY 241 FLDRDDGAKKVVYIETREPRERLLTAAHLLEVPAPHNDSATGEPPASSGPPSGGALG 300  
DB 241 FLDRDDGAKKVVYIETREPRERLLTAAHLLEVPAPHNDSATGEPPASSGPPSGGALG 300  
QY 301 PRALFASRVPGQRYVVAERDGRRLPAVAHSVTLSEEAAGAYAPLTAOGTILINRYL 360  
DB 301 PRALFASRVPGQRYVVAERDGRRLPAVAHSVTLSEEAAGAYAPLTAOGTILINRYL 360  
QY 361 ASCYAVIEESHNAHRAFAFRLAHLAALAPARDRGDSSGGDRGGGGVATLAPGA 420  
DB 361 ASCYAVIEESHNAHRAFAFRLAHLAALAPARDRGDSSGGDRGGGGVATLAPGA 420  
QY 421 ADAPGAGATAGIHWTYSQLYQIGTWLIDSEALHPLGMAYKSSXSRGAGGAREGA 475  
DB 421 ADAPGAGATAGIHWTYSQLYQIGTWLIDSEALHPLGMAYKSSXSRGAGGAREGA 475

# RESULT 6 US-09-325-256-22

Sequence 22, Application US/09325256  
Patent No. 6444793

## GENERAL INFORMATION:

APPLICANT: PEPIINSKY, R. BLAKE  
APPLICANT: BAKER, DARREN P.  
APPLICANT: WEN, DINGYI  
APPLICANT: WILLIAMS, KEVIN P.

APPLICANT: GARGER, ELLEN A.  
APPLICANT: TAYLOR, FREDERICK R.  
APPLICANT: GALDES, ALPHONSE  
APPLICANT: PORTER, JEFFREY

TITLE OF INVENTION: HYDROPHOBICALLY-MODIFIED PROTEIN COMPOSITIONS AND  
TITLE OF INVENTION: METHODS

FILE REFERENCE: BY-067.01

CURRENT APPLICATION NUMBER: US/09/325,256

PRIOR FILING DATE: 1999-06-03

PRIOR APPLICATION NUMBER: 60/099,800

PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/078,935

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/089,685

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/067,423

PRIOR FILING DATE: 1997-12-03

PRIOR APPLICATION NUMBER: PCT/US98/25676

PRIOR FILING DATE: 1998-12-03

NUMBER OF SEQ ID NOS: 31

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 22

LENGTH: 475

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Xaa may be any amino acid

US-09-325-256-22

Query Match 99.9%; Score 2467; DB 4; Length 475;  
Best Local Similarity 100.0%; Pred. No. 8.2e-239;

Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHKITPLAYKQFIPNVAEKTIGASG 60  
DB 1 MLLARCLLVVSSLLVCSGLACGPGFGKRRHKITPLAYKQFIPNVAEKTIGASG 60  
QY 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNQMPGV 120  
DB 61 RYEGKISRNSERFKELTPYNDIIFKDEENTGADRLMTORCKDKLNLAISVMNQMPGV 120  
QY 121 KLRVTEGMDEDDHSEESLHYEGRAVDITTSRDRSKYGMGLARLAVEAGFDVYVESKAH 180  
DB 121 KLRVTEGMDEDDHSEESLHYEGRAVDITTSRDRSKYGMGLARLAVEAGFDVYVESKAH 180  
QY 181 IHCYKAENSVAAKSGGCGPGSATVHLEGGTKLVKDLSPGDRVLAADOGRLLYSDFLT 240  
DB 181 IHCYKAENSVAAKSGGCGPGSATVHLEGGTKLVKDLSPGDRVLAADOGRLLYSDFLT 240



QY 181 IHCYKAEKNSVAAKSGCGFPGSATVHLEOGCTKLYKDLSPGDRVLAADDOGRLLYSDFLT 240  
DB 181 IHCYKAEKNSVAAKSGCGFPGSATVHLEOGCTKLYKDLSPGDRVLAADDOGRLLYSDFLT 240  
QY 241 FLDDDDAKKVFYIETREPERERLLTFAHLLFPAPHNDSATGEPKSSGSGPPSGALG 300  
DB 241 FLDDDDAKKVFYIETREPERERLLTFAHLLFPAPHNDSATGEPKSSGSGPPSGALG 300  
QY 301 PRALFASVRPGQVYVVAERDGRLLPAVHSVTLSEBAAGAYAPLTAGTILLINRYL 360  
DB 301 PRALFASVRPGQVYVVAERDGRLLPAVHSVTLSEBAAGAYAPLTAGTILLINRYL 360  
QY 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGSDGSGGGRGGRVALTAPGA 420  
DB 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGSDGSGGGRGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWYSOLLTOIGTWLDSBALHPLGMAVSSXSRGAGGARREGA 475  
DB 421 ADAPGAGATAGIHWYSOLLTOIGTWLDSBALHPLGMAVSSXSRGAGGARREGA 475

## RESULT 7

US-08-748-591-4  
; Sequence 4, Application US/08748591  
; Patent No. 5739811  
; GENERAL INFORMATION:  
; APPLICANT: Epstein, Ervin  
; APPLICANT: Hu, Zhilan  
; APPLICANT: Bonifas, Jeanette  
; TITLE OF INVENTION: Mutant Human Hedgehog Gene  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish and Richardson  
; STREET: 2200 Sand Hill Road  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/748,591  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sherwood, Pamela J  
; REGISTRATION NUMBER: 36,677  
; REFERENCE/DOCKET NUMBER: 06510/067001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 322-5070  
; TELEFAX: (415) 854-0875  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 462 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-748-591-4

Query Match 97.5%; Score 2407; DB 1; Length 462;  
Best Local Similarity 100.0%; Pred. No. 8.2e-233;

Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTARCLLVVSSLLVCSGLACGPGRGFKRRHPPKLPPLAYKQFIPNVAEKTGLASG 60  
DB 1 MLTARCLLVVSSLLVCSGLACGPGRGFKRRHPPKLPPLAYKQFIPNVAEKTGLASG 60  
QY 61 RYEEKISNSRERKELTPNYPDIIIFKDEENTGADRLMTORCKKLNALAISVNNQWPGV 120  
DB 61 RYEEKISNSRERKELTPNYPDIIIFKDEENTGADRLMTORCKKLNALAISVNNQWPGV 120

DB 61 RYEEKISNSRERKELTPNYPDIIIFKDEENTGADRLMTORCKKLNALAISVNNQWPGV 120  
QY 121 KLRTYEGMDDEGHSESESLYEGRANDITTSDDRSKYGMILARLAVBAGFDMWYESSKAH 180  
DB 121 KLRTYEGMDDEGHSESESLYEGRANDITTSDDRSKYGMILARLAVBAGFDMWYESSKAH 180  
QY 181 IHCYKAEKNSVAAKSGCGFPGSATVHLEOGCTKLYKDLSPGDRVLAADDOGRLLYSDFLT 240  
DB 181 IHCYKAEKNSVAAKSGCGFPGSATVHLEOGCTKLYKDLSPGDRVLAADDOGRLLYSDFLT 240  
QY 241 FLDDDDAKKVFYIETREPERERLLTFAHLLFPAPHNDSATGEPKSSGSGPPSGALG 300  
DB 241 FLDDDDAKKVFYIETREPERERLLTFAHLLFPAPHNDSATGEPKSSGSGPPSGALG 300  
QY 301 PRALFASVRPGQVYVVAERDGRLLPAVHSVTLSEBAAGAYAPLTAGTILLINRYL 360  
DB 301 PRALFASVRPGQVYVVAERDGRLLPAVHSVTLSEBAAGAYAPLTAGTILLINRYL 360  
QY 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGSDGSGGGRGGRVALTAPGA 420  
DB 361 ASCYAVIEESHMAHRAFPRLAHALLAALAPARTDRGSDGSGGGRGGRVALTAPGA 420  
QY 421 ADAPGAGATAGIHWYSOLLTOIGTWLDSBALHPLGMAVSS 462  
DB 421 ADAPGAGATAGIHWYSOLLTOIGTWLDSBALHPLGMAVSS 462

## RESULT 8

US-08-748-591-9  
; Sequence 9, Application US/08748591  
; Patent No. 5739811  
; GENERAL INFORMATION:  
; APPLICANT: Epstein, Ervin  
; APPLICANT: Hu, Zhilan  
; APPLICANT: Bonifas, Jeanette  
; TITLE OF INVENTION: Mutant Human Hedgehog Gene  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish and Richardson  
; STREET: 2200 Sand Hill Road  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/748,591  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sherwood, Pamela J  
; REGISTRATION NUMBER: 36,677  
; REFERENCE/DOCKET NUMBER: 06510/067001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 322-5070  
; TELEFAX: (415) 854-0875  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 462 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-748-591-9

Query Match 97.5%; Score 2407; DB 1; Length 462;  
Best Local Similarity 100.0%; Pred. No. 8.2e-233;

Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MILLARCLLVVSSLLVCSGLACGPGRGFRGRRHPPKLTPLAYKQFIPNVAEKTIGASG 60
DB 1 MILLARCLLVVSSLLVCSGLACGPGRGFRGRRHPPKLTPLAYKQFIPNVAEKTIGASG 60
OY 61 REEGKISRSEFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATISVMNQPGV 120
DB 61 REEGKISRSEFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATISVMNQPGV 120
OY 121 KLRVTEGMEDEGHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWVYESKAH 180
DB 121 KLRVTEGMEDEGHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWVYESKAH 180
OY 181 IHCSVKAENSVAAKSGGCPGSAVTHLEGGGKTKLYKDLSPGDRVLAADDQGRILYSDFLT 240
DB 181 IHCSVKAENSVAAKSGGCPGSAVTHLEGGGKTKLYKDLSPGDRVLAADDQGRILYSDFLT 240
OY 241 FLDRDGAKKVYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSPSGGALG 300
DB 241 FLDRDGAKKVYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSPSGGALG 300
OY 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360
DB 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360
OY 361 ASCYAVIEHSHVAHRAFAFRLAHLAALAPARTDRGDSGGGGRGGRVALTAPGA 420
DB 361 ASCYAVIEHSHVAHRAFAFRLAHLAALAPARTDRGDSGGGGRGGRVALTAPGA 420
OY 421 ADAPGAGTAGIHWYSQLLYOIGTWLDSSEALHPGLMAVKSS 462
DB 421 ADAPGAGTAGIHWYSQLLYOIGTWLDSSEALHPGLMAVKSS 462

```

```

; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; TOPOLOGY: 1linear
; MOLECULE TYPE: protein
US-08-946-329A-20

Query Match      83.8%; Score 2068; DB 3; Length 437;
Best Local Similarity 87.4%; Pred. No. 7.5e-199;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

OY 1 MILLARCLLVVSSLLVCSGLACGPGRGFRGRRHPPKLTPLAYKQFIPNVAEKTIGASG 60
DB 2 MILLARCLLVVSSLLVCSGLACGPGRGFRGRRHPPKLTPLAYKQFIPNVAEKTIGASG 61
OY 61 REEGKISRSEFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATISVMNQPGV 120
DB 62 REEGKISRSEFKELTPYNDIIFKDEENTGADRLMTORCKDKLNALATISVMNQPGV 121
OY 121 KLRVTEGMEDEGHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWVYESKAH 180
DB 122 KLRVTEGMEDEGHSESLHYEGRAVDITTSRDRSKYGLARLAVEAGFDWVYESKAH 181
OY 181 IHCSVKAENSVAAKSGGCPGSAVTHLEGGGKTKLYKDLSPGDRVLAADDQGRILYSDFLT 240
DB 182 IHCSVKAENSVAAKSGGCPGSAVTHLEGGGKTKLYKDLSPGDRVLAADDQGRILYSDFLT 241
OY 241 FLDRDGAKKVYVETREPRERLLTAAHLLFVAPHNDSATGEPEASSGSPSGGALG 300
DB 242 FLDRDGAKKVYVETREPRERLLTAAHLLFVAPHND-----SGPTPG 286
OY 301 PALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 360
DB 287 PALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAVAPLTAOGTILINRYL 346
OY 361 ASCYAVIEHSHVAHRAFAFRLAHLAALAPARTDRGDSGGGGRGGRVALTAPGA 420
DB 347 ASCYAVIEHSHVAHRAFAFRLAHLAALAPARTDRGDSGGGGRGGRVALTAPGA 395
OY 421 ADAPGAGTAGIHWYSQLLYOIGTWLDSSEALHPGLMAVKSS 462
DB 396 TEARGAEPAGIHWYSQLLYOIGTWLDSSEALHPGLMAVKSS 437

RESULT 10
US-08-567-357A-20
; Sequence 20, Application US/08567357A
; Patent No. 6132728
; GENERAL INFORMATION:
; APPLICANT: Beachy, Phillip A.
; APPLICANT: Moon, Randall T.
; APPLICANT: Porter, Jeffrey A.
; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; FILING DATE: 02-DEC-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/349,498
; REFERENCE/DOCKET NUMBER: 07265/140001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099

```

ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A. 38,347  
REGISTRATION NUMBER: 07265/080001  
REFERENCE/DOCKET NUMBER: 07265/080001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-567-357A-20

Query Match 83.8%; Score 2068; DB 4; Length 437;  
Best Local Similarity 87.4%; Pred. No. 7.5e-199;  
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLVSGLAGPGRGFGKRRHPPKLPPLAYKQIPVNAEKTLAGSG 60  
DB 2 LLLARCFVLVLAASSLVCPGLACGPGRGFGKRRHPPKLPPLAYKQIPVNAEKTLAGSG 61  
QY 61 RYEGKISRNSRFKELIPNYPNDIIFKDEENTGADRLMTORCKLNALAI SVNMOPGV 120  
DB 62 RYEGKITRNSRFKELIPNYPNDIIFKDEENTGADRLMTORCKLNALAI SVNMOPGV 121  
QY 121 KLRTGEGWDEGHHSESLAHYEGRAVDITTSDRDSKYGMRLAVEGFDWYYESKAH 180  
DB 122 KLRTGEGWDEGHHSESLAHYEGRAVDITTSDRDSKYGMRLAVEGFDWYYESKAH 181  
QY 181 IHCSVKAENSVAAKSGCGPFSATVHLEQGCTKLVKDLRPDRLAADDOGRLLYSDFLT 240  
DB 182 IHCSVKAENSVAAKSGCGPFSATVHLEQGCTKLVKDLRPDRLAADDOGRLLYSDFLT 241  
QY 241 FLDDDDAKKFFVYIERPERERLLTAAHLLFVAPHNDSATGPEASSGSGPPSGALG 300  
DB 242 FLDDDEGAKKFFVYIERPERERLLTAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASRVPRGQRYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAOGTILINRVL 360  
DB 287 PSALFASRVPRGQRYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAOGTILINRVL 346  
QY 361 ASCYAVIEEHSWAHRAFPFLAHLAALAPARTDRGDSGGDRGGGGRVALTPAGA 420  
DB 347 ASCYAVIEEHSWAHRAFPFLAHLAALAPARTD-----GGGGGSIIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWSQLLYOIGTWLDSALHPPLGMAVKS 462  
DB 396 TEARGAEPYTAGIHWSQLLYHIGTWLDSETHPLGMAVKS 437

RESULT 11  
US-08-729-743A-20  
Sequence 20, Application US/08729743A  
Patent No. 6214794  
GENERAL INFORMATION:  
APPLICANT: Beachy, Philip A.  
APPLICANT: Moon, Randall T.  
APPLICANT: Porter, Jeffrey A.  
TITLE OF INVENTION: NOVEL HEDGEGHOG-DERIVED POLYPEPTIDES  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95

SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/729,743A  
FILING DATE: 07-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/567,357  
FILING DATE: 04-DEC-1995  
APPLICATION NUMBER: 08/349,498  
FILING DATE: 02-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/099001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-729-743A-20

Query Match 83.8%; Score 2068; DB 4; Length 437;  
Best Local Similarity 87.4%; Pred. No. 7.5e-199;  
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

QY 1 MLLARCLLVVSSLVSGLAGPGRGFGKRRHPPKLPPLAYKQIPVNAEKTLAGSG 60  
DB 2 LLLARCFVLVLAASSLVCPGLACGPGRGFGKRRHPPKLPPLAYKQIPVNAEKTLAGSG 61  
QY 61 RYEGKISRNSRFKELIPNYPNDIIFKDEENTGADRLMTORCKLNALAI SVNMOPGV 120  
DB 62 RYEGKITRNSRFKELIPNYPNDIIFKDEENTGADRLMTORCKLNALAI SVNMOPGV 121  
QY 121 KLRTGEGWDEGHHSESLAHYEGRAVDITTSDRDSKYGMRLAVEGFDWYYESKAH 180  
DB 122 KLRTGEGWDEGHHSESLAHYEGRAVDITTSDRDSKYGMRLAVEGFDWYYESKAH 181  
QY 181 IHCSVKAENSVAAKSGCGPFSATVHLEQGCTKLVKDLRPDRLAADDOGRLLYSDFLT 240  
DB 182 IHCSVKAENSVAAKSGCGPFSATVHLEQGCTKLVKDLRPDRLAADDOGRLLYSDFLT 241  
QY 241 FLDDDDAKKFFVYIERPERERLLTAAHLLFVAPHNDSATGPEASSGSGPPSGALG 300  
DB 242 FLDDDEGAKKFFVYIERPERERLLTAAHLLFVAPHND-----SGPTPG 286  
QY 301 PRALFASRVPRGQRYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAOGTILINRVL 360  
DB 287 PSALFASRVPRGQRYVVAERDGRLLPAVHSVTLSEEAAGAYAPLTAOGTILINRVL 346  
QY 361 ASCYAVIEEHSWAHRAFPFLAHLAALAPARTDRGDSGGDRGGGGRVALTPAGA 420  
DB 347 ASCYAVIEEHSWAHRAFPFLAHLAALAPARTD-----GGGGGSIIP-AAQSA 395  
QY 421 ADAPGAGATAGIHWSQLLYOIGTWLDSALHPPLGMAVKS 462  
DB 396 TEARGAEPYTAGIHWSQLLYHIGTWLDSETHPLGMAVKS 437

RESULT 12  
US-09-057-860A-6  
Sequence 6, Application US/09057860A  
Patent No. 6277820  
GENERAL INFORMATION:  
APPLICANT: Arnon Rosenthal  
APPLICANT: Mary Hynes  
APPLICANT: Weilian Ye  
TITLE OF INVENTION: Method Of Dopaminergic And Serotonergic  
TITLE OF INVENTION: Neuron Formation From Neuroprogenitor Cells  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:

```

ADDRESS: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpacin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,860A
FILING DATE: 09-Apr-1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Syvobds, Craig G
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P1364
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-057-860A-6

```

```

: GENERAL INFORMATION:
: APPLICANT: The Johns Hopkins University School of Medicine
: TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: U.S.A.
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/15463
: FILING DATE: 01-DEC-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Halle, Lisa A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 07265/080W01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5070
: TELEFAX: 619/678-5099
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 437 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US95-15463-20

Query Match      83.8%; Score 2068; DB 5; Length 437;
Best Local Similarity 87.4%; Pred. No. 7.5e-199;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

Oy 1 MLLARCLLVVSLVCSGLACGPGRGFKRRPKLTPLAYKQFIPIVNAEKTIGASG 60
Db 2 LLLARCLVLVASSLVCPGLACGPGRGFKRRPKLTPLAYKQFIPIVNAEKTIGASG 61
Oy 61 RYEGKISNSRFKELTPNYNPDIIFKDENTGADRLTQCKDKLNLAISVNMOPGV 120
Db 62 RYEGKITNSRFKELTPNYNPDIIFKDENTGADRLTQCKDKLNLAISVNMOPGV 121
Oy 121 KLRVTEGDEGHHSEESLHTEGRAVDITTSDRDSKYGMLARLAVEGFPWVYYESKAH 180
Db 122 KLRVTEGDEGHHSEESLHTEGRAVDITTSDRDSKYGMLARLAVEGFPWVYYESKAH 181
Oy 181 IHCYSKAENSYAAKSGCGFPGSATVNHLEOGSTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240
Db 182 IHCYSKAENSYAAKSGCGFPGSATVNHLEOGSTKLVKDLSPGDRVLAADDOGRLLYSDFLT 241
Oy 241 FLDRDGAKKVYVYETREPERRLTLTAHLLEFVAPHNDSATGPEEASGSGPESGALG 300
Db 242 FLDRDGAKKVYVYETLEPERRLTLTAHLLEFVAPHNDSATGPEEASGSGPESGALG 300
Oy 301 PRALEFASVNRGORYVVAERDGRRLPAAVHSTVLSSEAGAYAPLTAOCTILINVL 360
Db 287 PSALFASVNRGORYVVAERDGRRLPAAVHSTVLSSEAGAYAPLTAOCTILINVL 360
Oy 361 ASCYAVIEHSHMAHRAFPRLAHLAALAPARTDRGDSGGGCGGGRVALTARGA 420
Db 347 ASCYAVIEHSHMAHRAFPRLAHLAALAPARTDRGDSGGGCGGGRVALTARGA 420
Oy 421 ADAPGATAGIHWSQLLYIGTWLDSSEALHPLGMAVKSS 462
Db 396 TEARGAEPTAGIHWSQLLYIGTWLDSSEALHPLGMAVKSS 437

RESULT 15
```

```

PCT-US95-15923-20
: Sequence 20, Application PC/US9515923
: GENERAL INFORMATION:
: APPLICANT: The Johns Hopkins University School of Medicine, et al.
: TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: U.S.A.
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/15923
: FILING DATE: 04-DEC-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Halle, Lisa A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 07265/043W01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5070
: TELEFAX: 619/678-5099
: INFORMATION FOR SEQ ID NO: 20:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 437 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US95-15923-20

Query Match      83.8%; Score 2068; DB 5; Length 437;
Best Local Similarity 87.4%; Pred. No. 7.5e-199;
Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

Oy 1 MLLARCLLVVSLVCSGLACGPGRGFKRRPKLTPLAYKQFIPIVNAEKTIGASG 60
Db 2 LLLARCLVLVASSLVCPGLACGPGRGFKRRPKLTPLAYKQFIPIVNAEKTIGASG 61
Oy 61 RYEGKISNSRFKELTPNYNPDIIFKDENTGADRLTQCKDKLNLAISVNMOPGV 120
Db 62 RYEGKITNSRFKELTPNYNPDIIFKDENTGADRLTQCKDKLNLAISVNMOPGV 121
Oy 121 KLRVTEGDEGHHSEESLHTEGRAVDITTSDRDSKYGMLARLAVEGFPWVYYESKAH 180
Db 122 KLRVTEGDEGHHSEESLHTEGRAVDITTSDRDSKYGMLARLAVEGFPWVYYESKAH 181
Oy 181 IHCYSKAENSYAAKSGCGFPGSATVNHLEOGSTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240
Db 182 IHCYSKAENSYAAKSGCGFPGSATVNHLEOGSTKLVKDLSPGDRVLAADDOGRLLYSDFLT 241
Oy 241 FLDRDGAKKVYVYETREPERRLTLTAHLLEFVAPHNDSATGPEEASGSGPESGALG 300
Db 242 FLDRDGAKKVYVYETLEPERRLTLTAHLLEFVAPHNDSATGPEEASGSGPESGALG 300
Oy 301 PRALEFASVNRGORYVVAERDGRRLPAAVHSTVLSSEAGAYAPLTAOCTILINVL 360
Db 287 PSALFASVNRGORYVVAERDGRRLPAAVHSTVLSSEAGAYAPLTAOCTILINVL 360
Oy 361 ASCYAVIEHSHMAHRAFPRLAHLAALAPARTDRGDSGGGCGGGRVALTARGA 420
Db 347 ASCYAVIEHSHMAHRAFPRLAHLAALAPARTDRGDSGGGCGGGRVALTARGA 420
Oy 421 ADAPGATAGIHWSQLLYIGTWLDSSEALHPLGMAVKSS 462
Db 396 TEARGAEPTAGIHWSQLLYIGTWLDSSEALHPLGMAVKSS 437
```

Thu Feb 20 10:43:53 2003

us-09-827-110a-15.ra1

Page 10

Search completed: February 20, 2003, 10:14:13  
Job time : 14.4346 secs

.....

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:08:18 ; Search time 15.632 Seconds  
(without alignments)  
2921.176 Million cell updates/sec

Title: US-09-827-110A-15  
Perfect score: 2469  
Sequence: 1 MLILARCLLVLVSSLLVCS.....GMAVSSXSXRGAGGAREGA 475

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2062	83.5	437	2 A49425	Sonic hedgehog pro
2	2060	83.4	437	2 B53193	hedgehog homolog v
3	1807	73.2	425	2 A49424	patterning protein
4	1506	61.0	418	2 A53193	hedgehog homolog v
5	1498	60.7	444	2 S56765	morphogen Xhh prec
6	1459	59.1	415	2 A49426	sonic hedgehog gen
7	1183	47.9	396	2 B49425	Desert hedgehog pr
8	1027.5	41.6	336	2 C49425	Indian hedgehog pr
9	890	36.0	471	2 A46400	segment polarity p
10	449	18.2	94	2 G02735	desert hedgehog -
11	159.5	6.5	615	2 T29550	hypothetical prote
12	149	6.0	1207	2 T23754	hypothetical prote
13	138.5	5.6	1226	2 T24045	hypothetical prote
14	135.5	5.5	484	2 T34504	hypothetical prote
15	134.5	5.4	481	2 T27665	hypothetical prote
16	129	5.2	1147	1 MMAX1B	myosin heavy chain
17	128	5.2	629	2 T19563	hypothetical prote
18	126.5	5.1	1021	2 T23252	hypothetical prote
19	123.5	5.0	2639	2 T31328	fibronin - Chinese
20	121	4.9	868	2 T22281	hypothetical prote
21	119	4.8	318	2 T35213	probable secreted
22	118.5	4.8	702	2 G01840	T-box protein 2 -
23	116	4.7	1329	2 E70917	hypothetical glyci
24	115.5	4.7	940	2 E87250	(protein-PII) urid
25	115	4.7	486	2 H87311	hypothetical glyci
26	114	4.6	741	2 G70917	hypothetical glyci
27	111	4.5	333	2 A39065	homeotic protein E
28	111	4.5	1298	1 EDBE75	Immediate-early pr
29	110	4.5	462	2 T35852	probable UDP-N-ace

30	107.5	4.4	701	1 S46458	transcription fact
31	107	4.3	775	1 EDBE11	Immediate-early pr
32	106.5	4.3	482	2 E70554	hypothetical prote
33	106	4.3	482	2 JC7583	basic helix-loop-h
34	105.5	4.3	490	2 J09084	phosphatidylinosit
35	105.5	4.3	841	2 C87331	ISCC2, transposase
36	105.5	4.3	1524	2 T14070	peptide synthetase
37	105.5	4.3	7463	2 T36248	CDA peptide synthe
38	105	4.3	460	2 T33110	hypothetical prote
39	104.5	4.2	510	2 A55207	glycerol-3-phospha
40	104	4.2	402	1 S25077	monensin polyketid
41	104	4.2	1063	2 A40253	acidic nuclear pro
42	104	4.2	3161	2 T30342	protein HMP1 - Ye
43	103.5	4.2	975	2 S33121	homeotic protein C
44	102.5	4.2	413	2 E70661	probable PE protei
45	102.5	4.2	479	2 F70573	hypothetical prote

ALIGNMENTS

RESULT 1

A49425  
Sonic hedgehog protein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999

C:Accession: A49425

R:Reheland, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; M

A:Title: Sonic hedgehog, a member of a family of putative signalling molecules, is imp

A:Reference number: A49425; M0ID:94094334; PMID:7916661

A:Accession: A49425

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-437 <ECH>

A:Cross-references: GB:X76290

A:Note: authors translation is shown for the codon TCC at position 436

C:Genetics:

A:Gene: Shh

C:Superfamily: sonic hedgehog protein

Query Match 83.5%; Score 2062; DB 2; Length 437;  
Best Local Similarity 87.0%; Pred. No. 3.6e-147;  
Matches 402; Conservative 10; Mismatches 24; Indels 26; Gaps 3;

QY	1	MLILARCLLVIVSLVSGIACGPGRGKRRPKPLPLAYKOTIPVAEKTIGASG	60
DB	2	LLILARCLLVIVSLVSGIACGPGRGKRRPKPLPLAYKOTIPVAEKTIGASG	61
QY	61	RYEKGISRNSEPFKELTPNPNPDIIFKDEBNTGADRLMTORCKDLNALAISVNMQPGV	120
DB	62	RYEKGITRNSEPFKELTPNPNPDIIFKDEBNTGADRLMTORCKDLNALAISVNMQPGV	121
QY	121	KLRVTEGWDGHHSESLHYEGRAVDITTSDRDSRYGMLARLAVAGFDWYESSKAH	180
DB	122	RLRVTGWDGHHSESLHYEGRAVDITTSDRDSRYGMLARLAVAGFDWYESSKAH	181
QY	181	IHCYVKAKNSVAAKSGCGPFSATVHLEGGTKYKRLSPGDRVLADDDGRLYSFELT	240
DB	182	IHCYVKAKNSVAAKSGCGPFSATVHLEGGTKYKRLSPGDRVLADDDGRLYSFELT	241
QY	241	FLRDDGAKKVFYIEIREFRERILLTAHLLEFVAPHNDSATGEPKASGSGPPSGCALG	300
DB	242	FLRDDGAKKVFYIEIREFRERILLTAHLLEFVAPHNDSATGEPKASGSGPPSGCALG	301
QY	301	PRALFASRVPRGQRYVVAERDGRLLPAVHSVTLSEBAAGVADLTAGCTLLINRVL	360
DB	287	PSALFASRVPRGQRYVVAERDGRLLPAVHSVTLSEBAAGVADLTAGCTLLINRVL	346
QY	361	ASCAVAVIEHSHMARAFAPFLAHALLAALAPATDGGSGGSDRGGGGGVAVLTAPGA	420
DB	347	ASCAVAVIEHSHMARAFAPFLAHALLAALAPATDGGSGGSDRGGGGGVAVLTAPGA	395

```

0y 421 ADAPGAGATGGINHYSQLYIGTWTLDSEALHPLGMAVSS 462
      :| | | | | | | | | | | | | | | | | | | |
Db 396 TEARGAEPTAGIHNSQLYHIGTWLDSETHPLGMAVKS 437

```

## RESULT 2

hedgehog homolog vhh-1 - rat  
 C.Species: *Rattus norvegicus* (Norway rat)  
 C.Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999  
 C.Accession: B53193  
 R:Roelink, H.; Augsburg, A.; Heemskerk, J.; Kozh, V.; Norlin, S.; Rulz 1 Altaba, A.;  
 Cell 76, 761-775, 1994  
 A.Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgehog  
 A.Reference number: A53193; MUID:94170375; PMID:8124714  
 A.Accession: B53193  
 A.Status: preliminary  
 A.Molecule type: mRNA  
 A.Residues: 1-437 <ROE>  
 A.Cross-references: GB:I27340; NID:g452122; PIDN:AAA20999.1; PID:g452123  
 C.Superfamily: sonic hedgehog protein

Query Match	83.48;	Score 20650;	DB 2;	Length 437;
Best Local Similarity	87.58;	Pred. No. 5,1e-147;		
Matches 405; Conservative	8;	Mismatches 22;	Indels 28;	Gaps 4

QY	1	MLLARCLLVLVSSLLVSSGLACGPGRRGKRRHPPKLLPPLVYKOPRIPVAEKTTGASG	60
		:	
Db	2	LLLLARCFVALLASSLLVCGLACGPGRRGKRRHPPKLLPPLVYKOPRIPVAEKTTGASG	61
QY	61	RYEKGISRNSERFELTPNPNPDIIIFDEBNTGADRLMTORCKDKLNALISVNMOPGV	120
Db	62	RYEKGIRNRSERFELTPNPNPDIIIFDEBNTGADRLMTORCKDKLNALISVNMOPGV	121
QY	121	KLRTYEGMDEGHHSEESIYEGRAVDITTSDDRSKTYGMLARLAVAGDWMYYESKAH	180
Db	122	KLRTYEGMDEGHHSEESIYEGRAVDITTSDDRSKTYGMLARLAVAGDWMYYESKAR	181
QY	181	IHCYKAKENSVYAAKSGGCFPGSATAVHLEOGGTRLYVKLSPGDRVLAADDGRLLYSPFLT	240
Db	182	IHCYKAKENSVYAAKSGGCFPGSATAVHLEOGGTRLYVKLSPGDRVLAADDGRLLYSPFLT	241
QY	241	FLUDDDCAKKVYYIETREPRERILLTAHLVYAPNDSATGEBPEASSGSGPPSGCALG	300
Db	242	FLUDDECAKKVYYIETREPRERILLTAHLVYAPND-----SGPTPG	286
QY	301	PRALFAFRVPGQRYVVAEBDGRRLPLPAVSVTSEENAGAYAPLTNOGTLINRYL	360
Db	287	PSPLFAFRVPGQRYVVAERGGDRLLPLPAVSVTIREDAQAYAPLTADGTLINRYL	346
QY	361	ASCYAVTIEESHMAHRAFPRLAHALLAALAPARTDRGGDGGGGRGGRVALTPA-G	419
		:	
Db	347	ASCYAVTIEESHMAHRAFPRLAHALLAALAPARTD-----GGGGG-SIPAPQS	394
QY	420	AADAPGAGATAGIHWYSOLLYOIGTWLLDSEALHPHGMAYKSS	462
		:	
Db	395	VAERAGGAPRAGIHWYSOLLYHIGTWLLDSEALHPHGMAYKSS	437

### RESULT 3

patterning protein sonic hedgehog precursor - chicken  
 C:Species: Gallus gallus (chicken)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 21-Jul-2000  
 C:Accession: A49494  
 R:Riddle, R.D.; Johnson, R.L.; Laufer, E.; Tablin, C.  
 Cell 75, 1401-1416, 1993  
 A:Title: Sonic hedgehog mediates the polarizing activity of the ZPA.  
 A:Reference number: A49424; MUID:94094333; PMID:8269518  
 A:Accession: A49424  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-425 <RID>

A:Cross-references: GB:IL28099; NID:g453526; PIDN:AAA72428.1; PID:g453527  
C:Superfamily: sonic hedgehog protein  
F:1-26/Domain: signal sequence #status predicted <SIG>

Query Match	73.2%;	Score 1807;	DB 2;	Length 425;
Best [local] siml ar tv	77.8%;	Pred NO	5e-12A.	

Best Local Similarity 77.88; Pred. No. 5e-128;  
Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6

```

0Y 1 MLRLRLLVSVSLVSSGACPGRGFGRRHRPKLTPLAYKOFIPNVAEKTIGASG 60
Db 4 MLRLRLVGFICALVSSGLTCGPGRGIGRRHRPKLTPLAYKOFIPNVAEKTIGASG 63
QY 61 RYEGKISNSERFEKLAPNPNPDIIFKDEMTGADRLMTORCKDKLNAIAISVNMOPGV 120
Db 64 RYEGKITNSRFEKLTPNPNPDIIFKDEEMTGADRLMTORCKDKLNAIAISVNMOPGV 123
QY 121 KLRTYEGDEGHHSESLHTEGRAVDITTSDBRSXKGMRLAFLAVEGFMWYIESAH 180
Db 124 KLRTYEGDEGHHSESLHTEGRAVDITTSDBRSKXIGMLARLAVEGFMWYIESAH 183
QY 181 IHC5VKAENSVAAKSGGCGPGSATVHLEQGGTKLVKIDPSGDRVLAADOGRLTYSFLT 240
Db 184 IHC5VKAENSVAAKSGGCGPGSATVHLEHGCKLVKIDPSGDRVLAADADRLLYSFLT 243
QY 241 FLDRBDGAKKVFYVYIETREPRERLLTNAHLLFAP-NHDS-ATGEPASSGSGPEPSGA 298
Db 244 FLDRDSSRKLFYVYIETROPARLLTNAHLLFAPAHQNSGAVG----STSG----- 292
QY 299 LGRPALFASRVPCGORYVVAVERDGRLLRPLAAVHSVLSLEAGAVAPLTAOGTILINR 358
Db 293 ---QALFASNVKPGORVYVLE--GGQOLPLASVHSVLSLREASGAVAPLTAOGTILINR 347
QY 359 VLACSYAVIEEHSNAHRFAFAPRLAHALLAALAPARTDGGDSGGGDRGGGGGRVALTPAR 418
Db 348 VLACSYAVIEEHSNAHRFAFAPRLAOGLLAAL-----CP 381
QY 419 GAADAPGAGATAGIHWYSOLYQICTMILDEALHPLGMAVKSS 462
Db 382 DGAIPATATTTGTHWYSRLLYRIGSWLDDDLAHLPLGMAVPAS 425

```

## RESULT 4

hedgehog homolog vhh-1 - zebra fish  
C:Species: Brachydanio rerio (zebra fish)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 10-Dec-1999  
C:Accession: A53193  
R:Roellink, H., Augsburg, A., Heemskerk, J., Korzh, V., Norlin, S., Ruiz i Altaba, A.  
Cell 76, 761-775, 1994  
A:Title: Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgehog  
A:Reference number: A53193; MUID:94170375; PMID:8124714  
A:Accession: A53193  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-418 <ROE>  
A:Cross-references: CB:127585; NID:g452159; PID:g452160  
C:Superfamily: sonic hedgehog protein

Query Match	61.08;	Score 1506;	DB 2;	Length 418;
-------------	--------	-------------	-------	-------------

[illegible]





A:Cross-references: GB:X76292; NID:q443941; PID:CAA53924.1; PID:q443942  
 C:Genetics:  
 A:Gene: Dhh  
 C:Superfamily: sonic hedgehog protein

Query Match 47.9%; Score 1183; DB 2; Length 396;  
 Best Local Similarity 54.5%; Pred. No. 3.3e-81;  
 Matches 244; Conservative 55; Mismatches 87; Indels 62; Gaps 10;

3 LLARCLLVSVSSLVCSGLACGPGRG-FGKRHRPK-LTPAYKQFIPNVAEKLIGASG 60  
 7 LILPLCL-----ALALASGSGPGRGVGRRRYRKQVLPVLTQFVSMERKLGASG 61  
 DB 1 RYEGATISNRSEFKLTPNYNDIIFKDEBTGADRLMTQRCCKDLALAISSVNMOWPGV 120  
 DB 61 PEGVNTGSEFRDLVNYNDIIFKDEBTGADRLMTQRCCKDLALAISSVNMOWPGV 121  
 DB 121 KLRTVEGDEGHHSEESIHYEGRVADITTSDRDRSKYGMRLAVEGDFWVYESKAH 180  
 DB 122 RLRTVEGDEGHHADSDLHYEGRVADITTSDRDRSKYGMRLAVEGDFWVYESRHH 181  
 DB 181 IHCYKAENSVAAKSGCGPGSATVLEGGGKTLYKDLSPGDRVLADDOGRILYSDFLT 240  
 DB 182 IHVSKADNSLAVRAGCGEPGATVRLASGERKGLRELHRCQWVLAADAAGVPTVILL 241  
 DB 241 FLDRDDAKKVFYVETREPRRLTLTAHLLEFVAPHNDATGEPAASSGSGP-PSGAL 299  
 DB 242 FLDRDLQRRASFVAVETERPRKLLTPMHLVFAA-----RGPAPAPGDF 286  
 DB 300 GPRAFFASRRVRCQVYVVAERDGRILLPAAVHVTLSSEPAAGVAPLTAQGITLINRV 359  
 DB 287 AP--VFARRLRAGDSVLA---PGSDALQPARAVARA-REEVAVGAPLTATAGTLLVNDV 339  
 DB 360 LASCYAVIEHSHMAHRAFAFRLAHLAALAPARTDRGDSGGDGGGGRVALTAPG 419  
 DB 340 LASCYAVIESHQMAHRAFAFRLAHLAALGALL-----PG 372  
 DB 420 AADAFAGATACIHWYSQLTOIGTWL 447  
 DB 373 GAVOP-----TGMWYSRLYLAEELM 395

## RESULT 8

C49425  
 Indian hedgehog protein - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 11-Jan-2000  
 C:Accession: C49425  
 R:Epichelard, Y.; Epstein, D.J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J.A.; McMahon, J.A.; Title: Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the development of the mouse embryo. *Cell* 75, 1417-1430, 1993  
 A:Title: Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the development of the mouse embryo. *Cell* 75, 1417-1430, 1993  
 A:Reference number: A49425; MUID:94094334; PMID:7916661  
 A:Accession: C49425  
 A:Molecule type: mRNA  
 A:Status: preliminary  
 A:Residues: 1-336 <ECHO>  
 A:Cross-references: GB:X76291  
 C:Superfamily: sonic hedgehog protein

Query Match 41.6%; Score 1027.5; DB 2; Length 336;  
 Best Local Similarity 54.6%; Pred. No. 1.3e-69;  
 Matches 218; Conservative 39; Mismatches 77; Indels 65; Gaps 8;

71 ERFKELPNYPDIIFKDEBTGADRLMTQRCCKDLALAISSVNMOWPGVLTAVTEGDE 130  
 DB 1 ERFKELPNYPDIIFKDEBTGADRLMTQRCCKDLALAISSVNMOWPGVLTAVTEGDE 60  
 DB 131 DGHSESESLYEGRAVDITTSDRDRSKYGMRLAVEGDFWVYESKAHICSVKANS 190  
 DB 61 DGHSESESLYEGRAVDITTSDRDRSKYGMRLAVEGDFWVYESKAHICSVKANS 120  
 DB 191 VAAKSGCGPGSATVLEGGGKTLYKDLSPGDRVLADDOGRILYSDFLTFLDRDDAKK 250

DB 121 AAKTGGCGPACQVRLNGEGRVALSAVPGDRLAMGEGTPTPSDYLIFLDRPNRLR 180  
 DB 251 VFYVETREPRRLTLTAHLLEFVAPHNDATGEPAASSGSGPFGSGALGPALPASRV 310  
 DB 181 AFQVETQDPPRRRLALTRHLEFIA---DNHT-EPAAHF-----RATASHVO 224  
 DB 311 PCQVRYVVAERDGRLLPAAVHVTLSSEPAAGVAPLTAQGITLINRVLASCVVIEBH 370  
 DB 225 PCQVRYVVAERDGRLLPAAVHVTLSSEPAAGVAPLTAQGITLINRVLASCVVIEBH 279  
 DB 371 SWAHAFAPFRLAHL--AALAPARTDRGDSGGDGGGGRVALTAPGADAGACAT 429  
 DB 280 HLAQLAFWPLRFLPSLAWGSWTPE----- 304  
 DB 430 AGIHWYSQLTOIGTWLSDSEALHPLGMAVKSXSGAG 468  
 DB 305 -GVHSYPMQVLYRLGRLLLESTFHPGLMS-----GAG 335

## RESULT 9

A46400  
 segment polarity protein hedgehog - fruit fly (Drosophila melanogaster)  
 N:Alternate names: hh protein  
 C:Species: Drosophila melanogaster  
 C:Date: 07-Apr-1994 #sequence\_revision 18-Nov-1994 #text\_change 10-Dec-1999  
 C:Accession: A46400; JN0501; A43480  
 R:Tabata, T.; Eaton, S.; Kornberg, T.B.  
 Genes Dev. 6, 2635-2645, 1992  
 A:Title: The Drosophila hedgehog gene is expressed specifically in posterior compartment  
 A:Reference number: A46400; MUID:94040725; PMID:1340474  
 A:Accession: A46400  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-471 <TAB>  
 A:Cross-references: GB:S66384; NID:q435848; PID:q435849  
 A:Experimental source: Oregon-R  
 A:Note: Sequence extracted from NCBI backbone (NCBIN:138996, NCBI:138997)  
 R:Tabata, T.; Eaton, S.; Kornberg, T.B.  
 Genes Dev. 6, 2635-2645, 1992  
 A:Title: Structure and expression of hedgehog, a Drosophila segment-polarity gene  
 A:Reference number: JN0501; MUID:93185922; PMID:8166882  
 A:Accession: JN0501  
 A:Molecule type: mRNA  
 A:Residues: 1-471 <TRAS>  
 A:Cross-references: GB:L05404  
 A:Note: It is uncertain whether Met-1 or Met-51 is the initiator  
 R:Lee, J.J.; von Kessler, D.P.; Parks, S.; Beachy, P.A.  
 Cell 71, 33-50, 1992  
 A:Title: Secretion and localized transcription suggest a role in positional signaling  
 A:Reference number: A43480; MUID:93008241; PMID:1394430  
 A:Accession: A43480  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-471 <LEB>  
 A:Cross-references: GB:L02793; NID:q157609; PID:q157610  
 A:Note: Sequence extracted from NCBI backbone (NCBI:115418)  
 C:Comment: This protein is required for cell-cell communication.  
 C:Genetics:  
 A:Gene: hh  
 A:Cross-references: Flybase:FBgn0004644  
 A:Introns: 160/3; 248/1  
 C:Superfamily: sonic hedgehog protein  
 C:Keywords: transmembrane protein  
 F:62-82/Domain: transmembrane #status predicted <TM>

Query Match 36.0%; Score 890; DB 2; Length 471;  
 Best Local Similarity 44.3%; Pred. No. 4.3e-59;  
 Matches 202; Conservative 65; Mismatches 137; Indels 52; Gaps 10;

DB 6 RCL-----LVVSSLVCSGLACGPGRGFGKRRAPKLTPLAKQFIPNVAEKLIG 57  
 DB 59 RCLSLVSLVALLIVLPVWSPAHSCGPGRGGLRHR-ARMILYPLVLYKOTIPNISEVTNS 117

[illegible]

```

RESULT 10
G02735
desert hedgehog - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 11-Jan-2000
C:Accession: G02735
R:Drummond, I.A.
submitted to the EMBL Data Library, June 1996
A:Reference number: H01643
A:Accession: G02735
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-94 <DRU>
A:Cross-references: EMBL:U59748; NID:g1401271; PID:g1401272
C:Genetics:
A:Gene: hDHH
C:Superfamily: sonic hedgehog protein

Query Match      18.2%; Score 449; DB 2; Length 94;
Best Local Similarity 84.0%; Pred. No. 7e-27;
Matches 79; Conservative 13; Mismatches 2; Indels 0; Gaps 0

Oy      84  IFKKEENTGADRLTQRCCKDLNALAISVMNQVEVKLRVTEGWEDGDGHSEESLIHYEG 143
         ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       1  IFKKEENSGARLRTERCKERVNALAIYMMNMPEGRILRVTEGWEDGDHNAODSLHYEG 60

Oy      144 RAVDTITSDRRSKYGMLARLAVEAGFDWYYES 177
         ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       61 RALDITTSDDRNKKYGLLARLAVEAGFDWYYIGS 94

RESULT 11
T29550
hypochelical protein ZK377.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T29550
R:Nhan, M.; Hawkins, J.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans coemlvd ZK377.
A:Reference number: Z20639
A:Accession: T29550
A>Status: preliminary; translated from GB/EMBL/DDBJ

```

A:Molecule type: DNA  
A:Residues: 1-615 <NHR>  
A:Cross-references: EMBL:U088183, PUDN:AA552656.1, GSPDB:GN000028, CESP:ZK377.1  
A:Experimental source: strain Bristol N2; clone ZK377  
A:Genetics:  
A:Gene: CESP:ZK377.1  
A:Map position: X  
A:Introns: 46/1; 70/3; 157/2; 176/1; 245/2; 314/3; 402/3; 499/2

Query Match	6.5%	Score 159.5	DB 2	Length 615
Best Local Similarity	28.3%	Pred. No. 0.00044		
Matches	52	Conservative	33	Mismatches 74; Indels 25; Gaps 7;
QY	198	CFPGSAVHLEGGGATKLYKDI	SPGGDYLAADDOGRLLYSDFLTFELDRDGAKEYVYIET	257
Db	416	CFSGMEVEVETD-GIKMKIKDLKIGKPLVSM-	DEAFVTYSPVIMFLHKRDEIEAENLIET	473
QY	256	REPRRLRLTAAILLVFAVPHNDSAFGEPEASSGSPRSGALGRPALAASVRRPGQRYV		317
Db	474	ANGHS-ILKTDNHLTYV--SDCPTRS-----	DEKLVAAKEVKMDCKIHV	514
QY	318	VAERDGDRLRLPAAYHSAVYLSSEAAAGAYAPLTAOCTILINRYLASCYAVIEEHSWANRAAF		377
Db	515	TTD---SNWVIKKKKYSKISKVYE-TGISPSPLSTGDIIVNRYRLASCHSNLAKSLQGRFF		570
QY	378	AFER 381		
Db	571	SLYK 574		

[illegible]

Db 1053 VAMRESKAEARKCEVLSDSE-SGE-----VIADEIVRGRTNVGIYSPMT 1100  
QY 350 AOGTILINRVLASCYAVIEHSMARAPAF 380  
Db 1101 VEGSLIVDGVLLSCFSHLSHS-AHKLFEDF 1130

## RESULT 13

T24045

hypotheical protein R08B4.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T24045

R:White, S.  
submitted to the EMBL Data Library, November 1995

A:Reference number: 219834

A:Accession: T24045

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1226 &lt;WIL&gt;

A:Cross-references: EMBL:Z68008; PIDN:CAA92000.1; GSPDB:GN00028; CESP:R08B4.1

A:Experimental source: clone R08B4

C:Genetics:

A:Gene: CESP:R08B4.1

A:Map position: X

A:Introns: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 839/3; 877/1; 960/3; 1

Query Match 5.6%; Score 138.5; DB 2; Length 1226;  
Best Local Similarity 24.4%; Pred. No. 0.041;

Matches 50; Conservative 38; Mismatches 68; Indels 49; Gaps 8;

QY 189 NSVAAKSGCEPPGSAVHLDEGGTKLVKLSPGDRVLAADOGRLYSDFLTDRDGA 248

Db 1012 SALVATGACFSIDTWV-TTFPGKRMQDIDIGVLTADLE-KYFPIPTLMIRPEK 1069

QY 249 KKVIVITRPRERLTLTAHLF-----VAHNDATGEPEASGSGSPS 295

Db 1070 VOEFITMT-EGKTLRTSRFMNRKSGSYPOYIKMLPD-----GE----- 1113

QY 296 GCGALPRALFASRVPGQRYVVAERDGRLLPAVHVSVTLSEAGAVAPLTAQGTIL 355

Db 1114 -----AIFSDLEVGCYVVLVYKGRQOKIETITRSV-----RTGISPLTNKXII 1161

QY 356 INRVLASCYAVIEHS-----MAH 374

Db 1162 VNDMLASCYSEIQNTLTQTFEMAY 1186

RESULT 14

T34504

hypotheical protein ZK1290.12 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T34504

R:Taich, A.  
submitted to the EMBL Data Library, July 1995

A:Description: The sequence of C. elegans cosmid ZK1290.

A:Reference number: 221535

A:Accession: T34504

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-484 &lt;TAI&gt;

A:Cross-references: EMBL:U21308; PIDN:AAB93321.1; GSPDB:GN00020; CESP:ZK1290.12

A:Experimental source: strain Bristol N2; clone ZK1290

C:Genetics:

A:Gene: CESP:ZK1290.12

A:Map position: 2

A:Introns: 66/3; 124/2; 181/2; 392/1

Query Match 5.5%; Score 135.5; DB 2; Length 484;  
Best Local Similarity 24.5%; Pred. No. 0.021;

Matches 46; Conservative 42; Mismatches 65; Indels 35; Gaps 8;

QY 191 VAAKSGCEPPGSAVHLDEGGTKLVKLSPGDRVLAADOGRLYSDFLTDRDGA 250  
Db 283 VASGVPACTGNSKV-KTPAGEKSMADLSVGDMMVTV-EGKMTYTVRAASMLRLPDTKA 340  
QY 251 VEVVETREPRERLTLTAHLFVAPPHNSATGEPEASGSGPPGCGALPRALFASRV 310  
Db 341 AFIKL-TTEOGAIDMTPOHFYTKA--NCVTEME-----LVYAEEDMT 380  
QY 311 PGQRYVVAERDGRLLPAVHVSVTLSEEA-----AGAVAPLTAQGTILINRVLASCYAV 366  
Db 381 IGDCLMV--KENEKLVN-----TTISEKSYFETGYAPAPTEGTGDLIVDVYASCHNV 431  
QY 367 IEHSHWAH 374  
Db 432 VKANTLSH 439

## RESULT 15

T27665

hypotheical protein ZK1037.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T27665

R:Baishan, V.  
submitted to the EMBL Data Library, October 1996

A:Reference number: 220401

A:Accession: T27665

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-481 &lt;WIL&gt;

A:Cross-references: EMBL:Z81142; PIDN:CA803509.1; GSPDB:GN00023; CESP:ZK1037.10

A:Experimental source: clone ZK1037

C:Genetics:

A:Gene: CESP:ZK1037.10

A:Map position: 5

A:Introns: 37/3; 75/3; 141/3; 173/2; 253/1; 298/3; 310/1; 371/2; 424/3

Query Match 5.4%; Score 134.5; DB 2; Length 481;  
Best Local Similarity 24.3%; Pred. No. 0.024;

Matches 45; Conservative 34; Mismatches 67; Indels 39; Gaps 6;

QY 198 CFPGSAVHLDEGGTKLVKLSPGDRVLAADOGRLYSDFLTLD-----RDSGAKRVF 252

Db 314 CFPNDVAVNVYEKAVKRMDLEIGDMVEALDENG-----EDITFLPVKXWLRDPEQBAE 368

QY 253 VVETREPRERLTLTAHLFVAPPHNSATGEPEASGSGPPGCGALPRALFASRVAPG 312

Db 369 FLERSLDNGETFTLEKHLVY-----TTECRONS-----SEKISWESISAGKVNAV 415

QY 313 QRVYVVAERDGRLLPAVHVSVTLSE-----ENAGAVAPLTAQGTILINRVLASCYA 365

Db 416 DQFYLAQSE-----ALKRYRLVELLDIKRVKKTOIVAPMISQGHLLVKNKITSCHS 466

QY 366 VIEEH 370

Db 467 EVDHH 471

Search completed: February 20, 2003, 10:13:37  
Job time : 18.632 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:07:32 ; Search time 14.5662 Seconds

(without alignments)  
1352.533 Million cell updates/sec

Title: US-09-827-110A-15

Perfect score: 2469  
Sequence: 1 MLILAKCLLVLVSSILVCS.....GNAYKSSKRGAGGAREGA 475

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	2407	97.5	462	1 SHH_HUMAN	015465 homo sapien
2	2068	83.8	437	1 SHH_MOUSE	062226 mus musculu
3	2060	83.4	437	1 SHH_RAT	063673 rattus norv
4	1807	73.2	425	1 SHH_CHICK	091035 gallus gall
5	1539.5	62.4	432	1 SHH_CYNPY	090385 cynops pyr
6	1506	61.0	418	1 SHH_BRARE	092008 brachydanio
7	1498	60.7	444	1 SHH_XENLA	092000 xenopus lae
8	1432	58.0	416	1 TWHH_BRARE	090419 brachydanio
9	1282.5	51.9	408	1 IHH_CHICK	098938 gallus gall
10	1266	51.3	411	1 IHH_HUMAN	014623 homo sapien
11	1262	51.1	411	1 IHH_MOUSE	097812 mus musculu
12	1248	50.5	409	1 IHH_XENLA	091612 xenopus lae
13	1213	49.1	412	1 IHH_BRARE	098862 brachydanio
14	1183	47.9	386	1 DHH_MOUSE	061488 mus musculu
15	1177	47.7	386	1 DHH_HUMAN	043333 homo sapien
16	1079.5	43.7	396	1 DHH1_XENLA	091610 xenopus lae
17	1077.5	43.6	398	1 DHH2_XENLA	091611 xenopus lae
18	890	36.0	471	1 HH_DROME	002936 drosophila
19	854	34.6	481	1 HH_DROHY	056674 drosophila
20	602	24.4	121	1 SHH_RASEL	079858 rasbora ele
21	599	24.3	121	1 SHH_CARAU	079691 carassius a
22	599	24.3	121	1 SHH_PUNTE	079850 punctius tet
23	598	24.2	121	1 SHH_RASHE	079864 rasbora het
24	598	24.2	121	1 SHH_RASPA	079869 rasbora pav
25	597	24.2	121	1 SHH_TANAL	079915 tanichthys
26	595	24.1	121	1 SHH_AMBCH	079682 amblypharyn
27	594	24.1	121	1 SHH_DANNA	013235 danio aff.
28	594	24.1	121	1 SHH_DANNA	013234 danio aff.
29	594	24.1	121	1 SHH_DANAT	013248 danio aff.
30	594	24.1	121	1 SHH_DANAT	013245 danio frank
31	594	24.1	121	1 SHH_DANKE	079707 danio kerri
32	594	24.1	121	1 SHH_DANPU	079719 danio pulch
33	594	24.1	121	1 SHH_DEVDE	013241 devatio dev

34	594	24.1	121	1 SHH_DEVMA	013247 devatio mal
35	594	24.1	121	1 SHH_DEVPA	013250 devatio pat
36	590	23.9	121	1 SHH_PUNCO	079838 punctius con
37	396	16.0	88	1 DHH_BRARE	079729 brachydanio
38	293	11.9	58	1 IHH_CARAU	079693 carassius a
39	293	11.9	58	1 IHH_DANAT	013240 danio aff.
40	293	11.9	58	1 IHH_DANKE	079711 danio kerri
41	293	11.9	58	1 IHH_DANPU	079719 danio pulch
42	293	11.9	58	1 IHH_DEVDE	013243 devatio dev
43	293	11.9	58	1 IHH_PUNTE	079850 punctius tet
44	293	11.9	58	1 IHH_RASEL	079860 rasbora ele
45	289	11.7	58	1 SHH_PSEPR	079839 pseudorasbo

## ALIGNMENTS

RESULT 1  
ID SHH\_HUMAN STANDARD; PRT; 462 AA.  
AC 015465;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Sonic hedgehog protein precursor (SHH) (HNG-1).  
GN SHH.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal lung;  
RX MEDLINE=96070431; PubMed=7590746;  
RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,  
RA Gastier J.M., Epstein D.J., Gilbert D.J., Copeland N.G., Seldman C.E.,  
RA Jenkins N.A., Seldman J.G., McMahon A.P., Tabin C.;  
RT "Cloning, expression, and chromosomal location of SHH and IHH: two  
RT human homologues of the Drosophila segment polarity gene hedgehog.";  
RL Genomics 28:44-51(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Tate G., Kishimoto K., Mitsuura T.;  
RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in  
RT human cancer cell lines and embryonic organs.";  
RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).  
RN [3]  
RP SEQUENCE OF 1-187 FROM N.A.  
RA Strong C., Graves T., Sutterer C., Ozersky P.;  
RL Submitted (MUG-1997) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 119-167 FROM N.A.  
RX MEDLINE=95236997; PubMed=7720571;  
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,  
RA Zhao R., Seidman M.F., Fallon J.F., Beachy P.A.;  
RT "Products, genetic linkage and limb patterning activity of a murine  
RT hedgehog gene.";  
RL Development 120:3339-3353(1994).  
RN [5]  
RP PALMITOYLATION.  
RX MEDLINE=98256339; PubMed=9593755;  
RA Peplinsky R.B., Zeng C., Wen D., Rayhorn P., Baker D.P., Williams K.P.,  
RA Bixler S.A., Ambrose C.M., Garber E.A., Matkowski K., Taylor F.R.,  
RA Wang E.A., Gales A.;  
RT "Identification of a palmitic acid-modified form of human Sonic  
RT hedgehog.";  
RL J. Biol. Chem. 273:14037-14045(1998).  
RN [6]  
RP VARIANTS HPE ARG-31; GLY-117 AND ARG-117.  
RX MEDLINE=97051937; PubMed=8896572;  
RA Roessler E., Belloni E., Gaudenz K., Jay P., Berta P., Scherer S.W.,  
RA Tsui L.-C., Muenke M.;  
RT "Mutations in the human Sonic hedgehog gene cause holoprosencephaly.";

RL Nat. Genet. 14:357-360(1996).  
 RP [17]  
 RP VARIANTS HPE ARG-31; GLY-117; ARG-117; GLU-224; THR-226 AND THR-383.  
 RX MEDLINE=96027056; PubMed=9302262;  
 RA Roessler E., Belloni E., Gaudenz K., Vargas F., Scherer S.W.,  
 RA Tsai L.-C., Muenke M.,  
 RT "Mutations in the C-terminal domain of Sonic hedgehog cause  
 RT holoprosencephaly.";  
 RL Hum. Mol. Genet. 6:1847-1853(1997).  
 RN [8]  
 RP VARIANTS HPE HIS-100; GLN-188 AND ASN-222.  
 RX MEDLINE=99371775; PubMed=10441331;  
 RA Odent S., Atti-Bitach T., Belayu M., Mathieu M., Aug J.,  
 RA Delezo de A.L., Gall J.Y., Le Marec B., Munnich A., David V.,  
 RA Vekemans M.,  
 RT "Expression of the sonic hedgehog (SHH) gene during early human  
 RT development and phenotypic expression of new mutations causing  
 RT holoprosencephaly.";  
 RL Hum. Mol. Genet. 8:1683-1689(1999).  
 RN [9]  
 RP VARIANTS HPE V-88; K-115; R-236; 263-R--A-269 DEL; D-290; A-424 AND  
 RP L-436.  
 RX MEDLINE=20025757; PubMed=10556296;  
 RA Nanni L., Ming J.E., Bocian M., Steinhaus K., Bianchi D.M.,  
 RA Die-Smulders C., Giannotti A., Imazumi K., Jones K.L., Campo M.D.,  
 RA Martin R.A., Melnick P., Pierpont M.E.M., Robin N.H., Young I.D.,  
 RA Roessler E., Muenke M.,  
 RT "The mutational spectrum of the sonic hedgehog gene in  
 RT holoprosencephaly: SHH mutations cause a significant proportion of  
 RT autosomal dominant holoprosencephaly.";  
 RL Hum. Mol. Genet. 8:2479-2488(1999).  
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN  
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION  
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE  
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER  
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A  
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED  
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE  
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE  
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH  
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD  
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS  
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN PERAL INTESTINE, LIVER, LUNG, AND  
 CC KIDNEY. NOT EXPRESSED IN ADULT TISSUES.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- DISEASE: DEFECTS IN SHH ARE THE CAUSE OF HOLOPROSENCEPHALY (HPE).  
 CC HPE IS A CLINICALLY VARIABLE AND GENETICALLY HETEROGENEOUS  
 CC MALFORMATION IN WHICH THE DEVELOPING FOREBRAIN FAILS TO CORRECTLY  
 CC SEPARATE INTO RIGHT AND LEFT HEMISPHERES. IN THE MOST SEVERE FORM  
 CC (ALOBAR HPE), THERE IS NO INTERHEMISPHERIC FISSURE, A SINGLE BRAIN  
 CC VENTRICLE IS PRESENT. HPE IS ASSOCIATED WITH SEVERAL DISTINCT  
 CC FACIES AND PHENOTYPIC VARIABILITY. IN THE MOST EXTREME CASES,  
 CC ANOPHTHALMIA OR CYCLOPIA IS EVIDENT ALONG WITH A CONGENITAL  
 CC ABSENCE OF THE MATURE NOSE. THE LESS SEVERE FORM FEATURES FACIAL  
 CC DYSMORPHIA CHARACTERIZED BY OCULAR HYPERTELORISM, DEFECTS OF THE  
 CC UPPER LIP AND/OR NOSE, AND ABSENCE OF THE OLFACTORY NERVES OR  
 CC CORPUS CALLOSUM. THE MAJORITY OF HPE CASES ARE APPARENTLY  
 CC SPORADIC, ALTHOUGH CLEAR EXAMPLES OF AUTOSOMAL DOMINANT (AD)  
 CC INHERITANCE HAVE BEEN DESCRIBED. INTERESTINGLY, UP TO 30% OF

CC OBLIGATE CARRIERS OF HPE GENE IN AD PEDIGREES ARE CLINICALLY  
 CC UNAFFECTED.  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -1- DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Haematol.;  
 CC WWW="http://www.intoblog.fr/services/chromosome/Genes/SHHID378.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 CC EMBL: L38518; AAA62179.1; -;  
 CC DR HSPB: AC002484; AAB67604.1; -;  
 CC DR HSPB: O62226; 1YH.  
 CC DR MIM: 600725; -;  
 CC DR MIM: 142945; -;  
 CC DR InterPro: IPR000320; HH\_signal.  
 CC DR InterPro: IPR001767; Hedgehog\_hnt.  
 CC DR InterPro: IPR003586; Hedgehog\_hntC.  
 CC DR InterPro: IPR003587; Hedgehog\_hntN.  
 CC DR InterPro: IPR002203; Intein.  
 CC DR InterPro: IPR001657; SonicHH.  
 CC Pfam: PF01079; Hnt; 1.  
 CC Pfam: PF01085; HH\_signal; 1.  
 CC PRINTS: PR00632; SONICHHOG.  
 CC PRODOM: PD003042; HH\_signal; 1.  
 CC SMART: SM00305; HntC; 1.  
 CC SMART: SM00306; HntN; 1.  
 CC PROSITE: PS50817; INTEIN\_NTER; 1.  
 CC Developmental protein: Autocatalytic cleavage: Hydrolyase; Protease;  
 CC Signal: Lipoprotein; Palmitate; Disease mutation: Holoprosencephaly.  
 CC SIGNAL 1 23  
 CC CHAIN 24 462  
 CC CHAIN 24 197  
 CC CHAIN 198 462  
 CC SITE 197 198  
 CC SITE 243 243  
 CC SITE 267 267  
 CC ACT\_SITE 270 270  
 CC BINDING 197 197  
 CC DOMAIN 407 411  
 CC LIPID 24 24  
 CC VARIANT 31 31  
 CC VARIANT 88 88  
 CC VARIANT 100 100  
 CC VARIANT 115 115  
 CC VARIANT 117 117  
 CC VARIANT 117 117  
 CC VARIANT 188 188  
 CC VARIANT 222 222  
 CC VARIANT 224 224  
 CC VARIANT 226 226  
 CC VARIANT 236 236  
 CC VARIANT 263 269

/FTId-VAR\_009163.  
 O -> H (IN HPE: SPORADIC).  
 /FTId-VAR\_009164.  
 N -> K (IN HPE: FAMILIAL).  
 /FTId-VAR\_009165.  
 W -> G (IN HPE).  
 /FTId-VAR\_003620.  
 W -> R (IN HPE).  
 /FTId-VAR\_003621.  
 E -> Q (IN HPE: FAMILIAL).  
 /FTId-VAR\_009166.  
 D -> N (IN HPE: FAMILIAL).  
 /FTId-VAR\_009167.  
 V -> E (IN HPE).  
 /FTId-VAR\_009168.  
 A -> T (IN HPE: FAMILIAL).  
 /FTId-VAR\_009169.  
 S -> R (IN HPE: FAMILIAL).  
 /FTId-VAR\_009170.  
 MISSING (IN HPE: SPORADIC).

Query Match 97.5%; Score 2407; DB 1; Length 462;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-170;  
 Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLLARCLLVLSLVSLVSGIACGPGRGKRRHKKLPLPAVKQPIPAVAKETTGASG 60  
 DB 1 MLLARCLLVLSLVSLVSGIACGPGRGKRRHKKLPLPAVKQPIPAVAKETTGASG 60  
 QY 61 RYEGKISRNSERKELTPNPNPIIFKDEKENTGADRLMTORCKDLKALISVNMOPGV 120  
 DB 61 RYEGKISRNSERKELTPNPNPIIFKDEKENTGADRLMTORCKDLKALISVNMOPGV 120  
 QY 121 KLVNTEGMDGDHHSSESLHYEGRAVDITTSDBRSKYGMARLAVAGEFDWYYESKAH 180  
 DB 121 KLVNTEGMDGDHHSSESLHYEGRAVDITTSDBRSKYGMARLAVAGEFDWYYESKAH 180  
 QY 181 IHCSVNAENSVAAKSGCGFCGFSATVHLEOGTKLVKDLSDGDRVLADDDGRLLYSDFLT 240  
 DB 181 IHCSVNAENSVAAKSGCGFCGFSATVHLEOGTKLVKDLSDGDRVLADDDGRLLYSDFLT 240  
 QY 241 FLDRDDGAKVFYIETREPERELLTAHLFLFAPHNDSATGEPASGSGPSGALG 300  
 DB 241 FLDRDDGAKVFYIETREPERELLTAHLFLFAPHNDSATGEPASGSGPSGALG 300  
 QY 301 PRALFASRVPRGQRYVVAERDGRLLPAAVHSVTLSEBAGAYAPLTAQGTLLINRVL 360  
 DB 301 PRALFASRVPRGQRYVVAERDGRLLPAAVHSVTLSEBAGAYAPLTAQGTLLINRVL 360  
 QY 361 ASCYAVTEESHMAHAPAPRLAHLAALAPARTDNGSGCGGCGGGRVALTPAGA 420  
 DB 361 ASCYAVTEESHMAHAPAPRLAHLAALAPARTDNGSGCGGCGGGRVALTPAGA 420  
 QY 421 ADAPGAGATGIMHYSQLYQIGTWLDSALPLPLGMAVSS 462  
 DB 421 ADAPGAGATGIMHYSQLYQIGTWLDSALPLPLGMAVSS 462

RESULT 2  
 ID SHH\_MOUSE STANDARD; PRT: 437 AA.  
 AC 062226;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, last sequence update)  
 DT 16-OCT-2001 (Rel. 40, last annotation update)  
 DE Sonic hedgehog protein precursor (SHH) (HMG-1).  
 GN SHH OR HHGL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SOURCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=94094334; Pubmed=7916661;  
 RA Echard Y., Epstein D.J., St Jacques B., Shen L., Mohler J.,  
 RA McMahon J.A., McMahon A.P.;  
 RT "Sonic hedgehog, a member of a family of putative signalling  
 RT molecules, is implicated in the regulation of CNS polarity.";  
 RL Cell 75:1417-1430(1993).  
 RN [2]  
 RP REVISION TO 122.  
 RC STRAIN=C57BL/6J;  
 RA McMahon A.P.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.  
 RX MEDLINE=95236997; Pubmed=7720571;  
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,  
 RA Zhao R., Seidman M.F., Fallon J.F., Beachy P.A.;  
 RT "Products, genetic linkage and limb patterning activity of a murine  
 RT hedgehog gene.";  
 RL Development 120:3339-3353(1994).

RA [4]  
 RP FUNCTION, AND AUTOPROTEOLYTIC CLEAVAGE.  
 RX MEDLINE=95236997; Pubmed=7720571;  
 RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,  
 RA Beachy P.A., Jessell T.M.;  
 RT "Floor plate and motor neuron induction by different concentrations of  
 RT the amino-terminal cleavage product of sonic hedgehog  
 RT autoproteolysis.";  
 RL Cell 81:445-455(1995).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 34-195.  
 RX MEDLINE=96069744; Pubmed=7477329;  
 RA Hall T.M.T., Porter J.A., Beachy P.A., Leahy D.J.;  
 RT "A potential catalytic site revealed by the 1.7-A crystal structure of  
 RT the amino-terminal signalling domain of sonic hedgehog.";  
 RL Nature 378:212-216(1995).  
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN  
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION  
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE  
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER  
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A  
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED  
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE  
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE  
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH  
 CC FLOOR PLATE-AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD  
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS  
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF EMBRYONIC TISSUES  
 CC INCLUDING THE NOTOCHORD, VENTRAL NEURAL TUBE, FLOOR PLATE, LUNG  
 CC BUD, ZONE OF POLARIZING ACTIVITY AND POSTERIOR DISTAL MESENCHYME  
 CC OF LIMBS. IN THE ADULT, EXPRESSED IN LUNG AND NEURAL RETINA.  
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE DURING GASTRULATION.  
 CC -1- INDUCTION: BY RETINOIC ACID.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY.  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X76290; CMA53922.1; -  
 DR PDB: 1VHH; 29-JAN-96.  
 DR MEROPS: C46.002; -  
 DR MGD: MGI:98297; Shh.  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hntc.  
 DR InterPro: IPR003586; Hedgehog\_hntc.  
 DR InterPro: IPR003587; Hedgehog\_hntc.  
 DR InterPro: IPR002203; Intein.  
 DR InterPro: IPR001657; SonichH.  
 DR Pfam: PF01079; Hntc; 1.  
 DR Pfam: PF01085; HH\_signal; 1.  
 DR PRINTS: PR00632; SONICHOG.  
 DR ProDom: PD003042; HH\_signal; 1.  
 DR SMART: SM00305; Hntc; 1.



DR SMART: SM00306; Hintn; 1.  
 DR PROSITE: PS50817; INTEN\_N\_TER; 1.  
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KM Signal; Lipoprotein; Palmitate; 3D-structure.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 437 SONIC HEDGEHOG PROTEIN.  
 FT CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.  
 FT CHAIN 199 437 SONIC HEDGEHOG PROTEIN C-PRODUCT.  
 FT SITE 198 199 CLEAVAGE (AUTO-) (BY SIMILARITY).  
 FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY  
 FT SITE 244 244 SIMILARITY).  
 FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY  
 FT ACT\_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY  
 FT BINDING 198 198 SIMILARITY).  
 FT DOMAIN 383 387 CHOLESTEROL (BY SIMILARITY).  
 FT LIPID 25 25 POLY-GLY.  
 FT SEQUENCE 437 AA; 47773 MW; DOB872F08E7860E CRC64;

Query Match 83.8%; Score 2068; DB 1; Length 437;  
 Best Local Similarity 87.4%; Pred. No. 2.3e-145;  
 Matches 404; Conservative 8; Mismatches 24; Indels 26; Gaps 3;

OY 1 MILLARCLLVVSSLLVCSGLACPGRGFGRRPKKLTPLATYQFIIPNVAEKLGLASG 60  
 DB 2 LLLARCLVLIASSLLVCPGLACPGRGFGRRPKKLTPLATYQFIIPNVAEKLGLASG 61  
 OY 61 RYEGKISNRSERFKELTPNYNDIIFKDEENTGADRLMTORCKDKLNLALISVNMQGV 120  
 DB 62 REGKITNSRERFKELTPNYNDIIFKDEENTGADRLMTORCKDKLNLALISVNMQGV 121  
 OY 121 KIRTEGDEDEGHSESLHREGRAVDITTSDRSKYGMRLARLAVEAGFDMVYESKAH 180  
 DB 122 KIRTEGDEDEGHSESLHREGRAVDITTSDRSKYGMRLARLAVEAGFDMVYESKAH 181  
 OY 181 IHCSVKANSVAAGSGGPGSATVHLHOGGKLVKDIISPDGRVLAADOGRLVSDPLT 240  
 DB 182 IHCSVKANSVAAGSGGPGSATVHLHOGGKLVKDIISPDGRVLAADOGRLVSDPLT 241  
 OY 241 FLDRDGAKKVYVETETEPRERLTLTAHLFLVPHNDSATGPEASSSGSPSGALG 300  
 DB 242 FLDRDGAKKVYVETETEPRERLTLTAHLFLVPHNDSATGPEASSSGSPSGALG 300  
 OY 301 PRALFASVRRGQRYVVAERDGRRLIPAAVHSVTLSSEAGAVAPLTAOGTILINRVL 360  
 DB 287 PSALFASVRRGQRYVVAERDGRRLIPAAVHSVTLSSEAGAVAPLTAOGTILINRVL 346  
 OY 361 ASCYAVIEHSHARAFAPFRLAHLAALAPARTDRGDSGGGRGGGRVALTAPCA 420  
 DB 347 ASCYAVIEHSHARAFAPFRLAHLAALAPARTDRGDSGGGRGGGRVALTAPCA 395  
 OY 421 ADAPGAGTAGIHWSQLLYOIGTWLDSSEALHPLGMAVKSS 462  
 DB 396 TEARGAEPTAGIHWSQLLYOIGTWLDSSEALHPLGMAVKSS 437

RESULT 3  
 SHH\_RAT STANDARD; PRT; 437 AA.  
 AC 063673;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sonic hedgehog protein precursor (SHH).  
 GN SHH OR VHH-1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NX NCBI\_TaxID=10116;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Sprague-Dawley; TISSUE=Embryonic floor plate;  
 RC

FX MEDLINE-94170375; PubMed-8124714;  
 RA Roelink H., Augsburg A., Heemskerk J., Korch V., Norlin S.,  
 RA Ruiz I., Altaba A., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,  
 RA Dodd J.;  
 RT "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog  
 RT of hedgehog expressed by the notochord.";  
 RL Cell 76:761-775(1994).  
 CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN  
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION  
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESSSES THE  
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER  
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A  
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED  
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE  
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE  
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH  
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD  
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS  
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE NODE, NOTOCHORD, FLOOR PLATE,  
 CC AND POSTERIOR LIMB BUD MESENCHYME.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L27340; AAA20999.1; -  
 CC HSSP: O62226; 1VHH.  
 DR MEROPS: C46.002; -  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hnt.  
 DR InterPro: IPR003586; Hedgehog\_hntC.  
 DR InterPro: IPR003587; Hedgehog\_hntN.  
 DR InterPro: IPR002203; Intein.  
 DR InterPro: IPR001657; SonichH.  
 DR Pfam: PF01079; Hntc; 1.  
 DR Pfam: PF01085; HH\_signal; 1.  
 DR PRINTS: PR00632; SONICHOG.  
 DR PRODOM: PD003042; HH\_signal; 1.  
 DR SMART: SM00305; Hntc; 1.  
 DR SMART: SM00306; Hntn; 1.  
 DR PROSITE: PS50817; INTEN\_N\_TER; 1.  
 DR Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KM Signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 437 SONIC HEDGEHOG PROTEIN.  
 FT CHAIN 25 198 SONIC HEDGEHOG PROTEIN N-PRODUCT.  
 FT CHAIN 199 437 SONIC HEDGEHOG PROTEIN C-PRODUCT.  
 FT SITE 198 199 CLEAVAGE (AUTO-).  
 FT SITE 244 244 INVOLVED IN CHOLESTEROL TRANSFER (BY  
 FT SITE 244 244 SIMILARITY).  
 FT SITE 268 268 INVOLVED IN AUTO-CLEAVAGE (BY  
 FT ACT\_SITE 271 271 ESSENTIAL FOR AUTO-CLEAVAGE (BY



FT BINDING 198 198 SIMILARITY).  
 FT DOMAIN 383 387 CHOLESTEROL (BY SIMILARITY).  
 FT LIPID 25 25 POLY-GLY.  
 SQ SEQUENCE 437 AA; 47630 MW; 0DBEC19F0D1662A0 CRC64;  
 Query Match 83.4%; Score 2060; DB 1; Length 437;  
 Best Local Similarity 87.5%; Pred. No. 9.2e-145;  
 Matches 405; Conservative 8; Mismatches 22; Indels 28; Gaps 4;

QY 1 MLLAACLILVYSSILVCSGLACGRGRGRHKKLLPLAYKQFIPVNAKTTGASG 60  
 DB 2 LLLACFLVALLASLLVCPGLACGPRGRGKQHPKLLPLAYKQFIPVNAKTTGASG 61  
 QY 61 RYEGKISRNSERKELTPNNPDIFKEDENTGADRLMTORCKDKLALASVNMQPGV 120  
 DB 62 RYEGKITRNSERKELTPNNPDIFKEDENTGADRLMTORCKDKLALASVNMQPGV 121  
 QY 121 KLRVTEGMDGHHSESLHYEGRAVDITTSDBDRSKYGMARLAVEAGFDWVYESKAH 180  
 DB 122 KLRVTEGMDGHHSESLHYEGRAVDITTSDBDRSKYGMARLAVEAGFDWVYESKAH 181  
 QY 181 IHCSVAENSVAAKSGCGFCGSAVHLFEGGTLYVNDLSFGDRVLAADDGRLYSDFLT 240  
 DB 182 IHCSVAENSVAAKSGCGFCGSAVHLFEGGTLYVNDLSFGDRVLAADDGRLYSDFLT 241  
 QY 241 FLDRDQAKVYFYIETREPERELLTAHLLEFVAPHNDATGEPBASSGSPSGALG 300  
 DB 242 FLDRDQAKVYFYIETREPERELLTAHLLEFVAPHNDATGEPBASSGSPSGALG 300  
 QY 301 PRLAFASVRPGQRYVVAERDRLRLPAVHVSVTLSSEBAGAVAPLTAGGTILNRYL 360  
 DB 287 PSPLFASVRPGQRYVVAERDRLRLPAVHVSVTLSSEBAGAVAPLTAGGTILNRYL 346  
 QY 361 ASCYAVEESNMRAAPRLLAHLAALAPARTDGRSGSGGDRGGGGRVALTRP-G 419  
 DB 347 ASCYAVEESNMRAAPRLLAHLAALAPARTDGRSGSGGDRGGGGRVALTRP-G 419  
 QY 420 AADAPGAGATAGIHWSQTLQIGTWLIDSEALHPLGMAYKSS 462  
 DB 395 VAEARNGAPPAIGHWSQTLQIGTWLIDSEALHPLGMAYKSS 437

RESULT 4  
 SHH\_CHICK STANDARD; PRT; 425 AA.  
 AC 091035;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sonic hedgehog protein precursor (SHH).  
 GN SHH.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN 1;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=limb bud;  
 RX MEDLINE=94094333; PubMed=8269518;  
 RA Riddle R.D., Johnson R.L., Lauffer E., Tablin C.,  
 RA "Sonic hedgehog mediates the polarizing activity of the ZPA."  
 RL Cell 75:1401-1416(1993).  
 RN 12;  
 RP FUNCTION AND AUTOPROTEOLYTIC CLEAVAGE.  
 RX MEDLINE=95254654; PubMed=7736596;  
 RA Roelink H., Porter J.A., Chiang C., Tanabe Y., Chang D.T.,  
 RA Beachy P.A., Jessell T.M.;  
 RT "Floor plate and motor neuron induction by different concentrations of  
 RT the amino-terminal cleavage product of sonic hedgehog  
 RT autoproteolysis".  
 RL Cell 81:445-455(1995).

CC -1- FUNCTION: BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN  
 CC ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE TRANSCRIPTION  
 CC OF TARGET GENES. IN THE ABSENCE OF SHH, PTC REPRESENTS THE  
 CC CONSTITUTIVE SIGNALING ACTIVITY OF SMO. ALSO REGULATES ANOTHER  
 CC TARGET, THE GLI ONCOGENE. INTERCELLULAR SIGNAL ESSENTIAL FOR A  
 CC VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT: SIGNAL PRODUCED  
 CC BY THE NOTOCHORD THAT INDUCES VENTRAL CELL FATE IN THE NEURAL TUBE  
 CC AND SOMITES, AND THE POLARIZING SIGNAL FOR PATTERNING OF THE  
 CC ANTERIOR-POSTERIOR AXIS OF THE DEVELOPING LIMB BUD. DISPLAYS BOTH  
 CC FLOOR PLATE- AND MOTOR NEURON-INDUCING ACTIVITY. THE THRESHOLD  
 CC CONCENTRATION OF N-PRODUCT REQUIRED FOR MOTOR NEURON INDUCTION IS  
 CC 5-FOLD LOWER THAN THAT REQUIRED FOR FLOOR PLATE INDUCTION.  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE POSTERIOR LIMB BUD  
 CC MESENCHYME, THE HENSEN'S NODE, THE NOTOCHORD, AND THE FLOOR PLATE  
 CC OF THE NEURAL TUBE.  
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE AT STAGE 17 DURING THE  
 CC INITIATION OF LIMB BUD FORMATION. FROM THAT POINT ONWARDS, THE  
 CC EXPRESSION PATTERN EXACTLY MATCHES THE LOCATION OF THE ZONE OF  
 CC POLARIZING ACTIVITY (ZPA).  
 CC -1- INDUCTION: BY RETINOID ACID.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY.  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC  
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION  
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -  
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS  
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY  
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL  
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.isb-sib.ch/announce/>  
 CC OR SEND AN EMAIL TO [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: L28099; AAA72428.1; -  
 CC HSSP: O62226; 1VHH.  
 CC DR MEROPS: C46.001; -  
 CC DR InterPro: IPR000320; HH\_signal.  
 CC DR InterPro: IPR001767; Hedgehog\_hint.  
 CC DR InterPro: IPR003586; Hedgehog\_hintc.  
 CC DR InterPro: IPR003587; Hedgehog\_hintn.  
 CC DR InterPro: IPR002203; Intein.  
 CC DR InterPro: IPR001657; SonICHH.  
 CC Pfam: PF01079; Hintc; 1.  
 CC Pfam: PF01085; HH\_signal; 1.  
 CC PRINTS: PR00632; SONICHOG.  
 CC PRODOM: PD003042; HH\_signal; 1.  
 CC SMART: SM00305; Hintc; 1.  
 CC SMART: SM00306; Hintn; 1.  
 CC DR PROSITE: PS50817; INTEIN\_N\_TER; 1.  
 CC  
 CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 CC Signal; Lipoprotein; Palmitate.  
 CC FT SIGNAL 1 26  
 CC FT CHAIN 27 425  
 CC FT CHAIN 27 200  
 CC FT CHAIN 201 425  
 CC FT SITE 200 201  
 CC FT SITE 246 246  
 CC FT SITE 246 246  
 CC FT SITE 270 270  
 CC FT ACT\_SITE 273 273  
 CC FT BINDING 200 200  
 CC  
 CC SONIC HEDGEHOG PROTEIN.  
 CC SONIC HEDGEHOG PROTEIN N-PRODUCT.  
 CC SONIC HEDGEHOG PROTEIN C-PRODUCT.  
 CC CLEAVAGE (AUTO-) (PROBABLY).  
 CC INVOLVED IN CHOLESTEROL TRANSFER (BY  
 CC SIMILARITY).  
 CC INVOLVED IN AUTO-CLEAVAGE (BY  
 CC SIMILARITY).  
 CC ESSENTIAL FOR AUTO-CLEAVAGE (BY  
 CC SIMILARITY).  
 CC CHOLESTEROL (BY SIMILARITY).

FT DOMAIN 390 393 POLY-THR.  
 FT LIPID 27 27 PALMITATE (BY SIMILARITY).  
 SQ SEQUENCE 425 AA; 46474 MM; DA9627443DA0173 CRC64;

Query Match 73.2%; Score 1807; DB 1; Length 425;  
 Best Local Similarity 77.8%; Pred. No. 4,1e-126;  
 Matches 361; Conservative 22; Mismatches 37; Indels 44; Gaps 6;

CC 1 MILLARCLLLVSSLLVCSGLACGPGRGFRKRRPKLTPLAYKQFIPIVNAEKTIGASG 60  
 CC TERLEAGDEGHSESLHTEGRAVDITTSDRSKYGMARLAVEGFPMVYESKAH 180  
 CC 4 MILLRRLVLFICALVSSGLTGPGRGICKRRRPKLTPLAYKQFIPIVNAEKTIGASG 63  
 CC KLRTEGMDGHHSESLHTEGRAVDITTSDRSKYGMARLAVEGFPMVYESKAH 183  
 CC 181 IHCYKANSVAASGCGFPGSATVHLDEGGTKLVKDLSPGDRVLAADGRLYSDFLT 240  
 CC IHCYKANSVAASGCGFPGSATVHLDEGGTKLVKDLSPGDRVLAADGRLYSDFLT 243  
 CC 241 FLDRDDGAKKPYVETREPRERLITAAHLFPVAP-HNDS-ANGEPPASSGSGPPSGCA 298  
 CC 244 FLDRDDGAKKPYVETREPRERLITAAHLFPVAP-HNDS-ANGEPPASSGSGPPSGCA 292  
 CC 299 LGPRALFASRVPRGORYVVAERDGRRLPAAVHSVTLSSEAGAVAPLTAQGITLINR 358  
 CC 293 ---QALFASNVKPGQRYVLE--GGQOLPLPSVHSLREASGAVAPLTAQGITLINR 347  
 CC 359 VLASCVAVIEHSMHRAFAFRLAHALLAALAPARTDRGDSGGGGRVALTAP 418  
 CC 348 VLASCVAVIEHSMHRAFAFRLAHALLAALAPARTDRGDSGGGGRVALTAP 381  
 CC 419 GAADAPGAGATAGTWSQLYQIGTWLSDSEALHPICMAVKS 462  
 CC 382 DGAIFTAATTTTGHWSRLYRIGSWLDGDLHPICMAVAPAS 425

RESULT 5  
 ID SHH\_CYNPY STANDARD; PRT; 432 AA.  
 AC 090385;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sonic hedgehog protein precursor (SHH).  
 GN SHH.  
 OS Cynops pyrrhogaster (Japanese common newt).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Caudata; Salamandridae; Cynops.  
 OX NCBI\_TaxID=8330;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Embryo;  
 RX MEDLINE=96136334; PubMed=8573168;  
 RA Takahata T., Takahashi T.C., Inoue K., Ogawa M., Takashima K.;  
 RT "Activation of two Cynops genes, fork head and sonic hedgehog, in  
 RT animal cap explants.";  
 RL Biochem. Biophys. Res. Commun. 218:395-401(1996).  
 CC 1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED  
 CC IN LIMB FORMATION. PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND  
 CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND  
 CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH  
 CC FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE  
 CC TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH, PTC  
 CC REPRESENTS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE

CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- INDUCTION: ACTIVATED BY ACTININ, BASIC FIBROBLAST GROWTH FACTOR  
 CC (BFGF) AND FORK HEAD.  
 CC -1- PRT: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION  
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -  
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS  
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY  
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL  
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D63339; BAA09657.1; .  
 CC HSSP; Q62226; IVH.  
 CC MEROPS; C46.002; .  
 CC InterPro; IPR000320; HH\_signal.  
 CC InterPro; IPR001767; Hedgehog\_hntc.  
 CC InterPro; IPR003586; Hedgehog\_hntc.  
 CC InterPro; IPR003587; Hedgehog\_hntcn.  
 CC InterPro; IPR002203; Intein.  
 CC InterPro; IPR001657; SonicHH.  
 CC Pfam; PF01079; Hntc\_1.  
 CC Pfam; PF01085; HH\_signal; 1.  
 CC PRINTS; PR00632; SONICHHOG.  
 CC ProDom; PD003042; HH\_signal; 1.  
 CC SMART; SM00305; Hntc\_1.  
 CC SMART; SM00306; Hntcn\_1.  
 CC PROSITE; PS50817; INTEIN\_N\_TER; 1.  
 CC Developmental protein: Autocatalytic cleavage; Hydrolyase; Protease;  
 CC Signal; Lipoprotein; Palmitate.  
 CC FT SIGNAL 1 26  
 CC FT CHAIN 27 432  
 CC FT CHAIN 27 200  
 CC FT CHAIN 201 432  
 CC FT SITE 200 201  
 CC FT SITE 268 268  
 CC FT ACT\_SITE 271 271  
 CC FT BINDING 200 200  
 CC FT LIPID 27 27  
 CC FT SEQUENCE 432 AA; 47847 MM; B455C7E746C8E5A8 CRC64;

Query Match 62.4%; Score 1539.5; DB 1; Length 432;  
 Best Local Similarity 67.7%; Pred. No. 2.3e-106;  
 Matches 315; Conservative 36; Mismatches 75; Indels 39; Gaps 8;

CC 1 MILLARCLLLVSSLLVCSGLACGPGRGFRKRRPKLTPLAYKQFIPIVNAEKTIGASG 60  
 CC TERLEAGDEGHSESLHTEGRAVDITTSDRSKYGMARLAVEGFPMVYESKAH 180  
 CC 4 MILLRRLVLFICALVSSGLTGPGRGICKRRRPKLTPLAYKQFIPIVNAEKTIGASG 63  
 CC KLRTEGMDGHHSESLHTEGRAVDITTSDRSKYGMARLAVEGFPMVYESKAH 183  
 CC 181 IHCYKANSVAASGCGFPGSATVHLDEGGTKLVKDLSPGDRVLAADGRLYSDFLT 240  
 CC IHCYKANSVAASGCGFPGSATVHLDEGGTKLVKDLSPGDRVLAADGRLYSDFLT 243

```

OY 241 FLDRODGAKKVFYIETREPRELLTLTAALLFVA---PHNDSATGEPPASSGSGPPSGG 297
DB 244 FMDKEETVAKVFYIET--SREVRLLTAHLLEFVGAHFGND-----SGGDFRSVFG 293
OY 298 ALGPRALFASRRVPGRRVYVVAERODRLLPAAVHSVYLSEAAAGAYAPLTAGSTILLN 357
DB 294 SAGFRSMFASVAGHRRVLTV-DREG-RGLREATVRYVL-EERATGAYAPVTAHGTVID 350
OY 358 RVLASCYAVIEEHSMAHRAFPAPRLAHLAALAPARTRGDGGSGGGGGRVALTA 417
DB 351 RVLASCYAVIEEHSMAHRAFPAPRLVGLISLFFSP--QDYSHS----- 392
OY 418 PGAAADPGAGATAGIHWSQLLYQIGTWLIDSEALHPILGMAVKS 462
DB 393 -----PPAPSGEGVHMVSEILYIGTWTWLOEDTIIHPLGMAKSS 432

RESULT 6
SHH_BRARE
ID SHH_BRARE STANDARD: PRT: 418 AA.
AC 092008: 013170: 013171:
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sonic hedgehog protein precursor (SHH) (VHH-1).
GN SHH OR VHH1.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo:
RX MEDLINE-94170375; PubMed-8124714;
RA Koelink H., Augsburg A., Heemskerk J., Korzh V., Norlin S.,
RA Doud J., Tanabe Y., Placzek M., Edlund T., Jessell T.M.,
RA Ruzi J.;
RT "Floor plate and motor neuron induction by vhh-1, a vertebrate homolog
of hedgehog expressed by the notochord.";
RL Cell 76:761-775(1994).
[2]
RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.
RX MEDLINE-96014264; PubMed-7583153;
RA Ekker S.C., Ungar A.R., Greenstein P., von Kessler D.P., Porter J.A.,
RA Moon R.T., Beachy P.A.;
RT "Patterning activities of vertebrate hedgehog proteins in the
developing eye and brain.";
RL Curr. Biol. 5:944-955(1995).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE-96083328; PubMed-7579523;
RA Fietz M.J., Concordet J.-P., Barbosa R., Johnson R., Krauss S.,
RA McMahon A.P., Tabin C., Ingham P.W.;
RT "The hedgehog gene family in Drosophila and vertebrate development.";
RL Development Suppl. 43-51(1994).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE-99225458; PubMed-10207136;
RA Muller F., Chang B., Albert S., Fischer N., Tora L., Strahle U.;
RT "Intronic enhancers control expression of zebrafish sonic hedgehog in
floor plate and notochord.";
RL Development 126:2103-2116(1999).
[5]
RP SEQUENCE OF 30-92 AND 113-170 FROM N.A.
RC TISSUE=Muscle:
RX MEDLINE-97075114; PubMed-8917540;
RA Zarolaya R., Abouheif E., Meyer A.;
RT "Evolutionary analyses of hedgehog and Hoxd-10 genes in fish species
closely related to the zebrafish.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:13036-13041(1996).
[6]
RP FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF

```

```

CC CC PATTERNING EVENTS DURING DEVELOPMENT. SIGNAL PRODUCED BY THE
CC CC NOTOCHORD THAT INDUCES SOMITE PATTERNING, DORSO-VENTRAL PATTERNING
CC CC OF THE BRAIN AND EARLY PATTERNING OF THE DEVELOPING EYES. DISPLAYS
CC CC FLOOR PLATE-INDUCING ACTIVITY. BINDS TO THE PATCHED (PTC)
CC CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO
CC CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IN THE ABSENCE OF SHH,
CC CC PTC REPRESSES THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY
CC CC SIMILARITY).
CC CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE
CC CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE
CC CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM
CC CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).
CC CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRAL MIDLINE OF THE NEURAL
CC CC TUBE AND BRAIN. ALSO FOUND IN THE NOTOCHORD AND IN DEVELOPING FIN
CC CC BUD. IN THE DEVELOPING BRAIN, EXPRESSION OCCURS IN DOMAINS THAT
CC CC INCLUDE A DISCRETE REGION IN THE FLOOR OF THE DIENCEPHALON.
CC CC -1- DEVELOPMENTAL STAGE: FIRST DETECTABLE IN THE INNER CELL LAYER OF
CC CC THE EMBRYONIC SHIELD DURING GASTRULATION. BY 9.5 HRS OF
CC CC DEVELOPMENT, EXPRESSED IN A CONTINUOUS BAND THAT EXTENDS FROM THE
CC CC TAIL TO THE HEAD, THE ANTERIOR BOUNDARY OF EXPRESSION BEING
CC CC POSITIONED IN THE CENTER OF THE ANIMAL POLE ANTERIOR TO THE
CC CC PRESUMPTIVE MIDBRAIN.
CC CC -1- PTH: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY
CC CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN
CC CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF
CC CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-
CC CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS
CC CC TO PLAY AN ESSENTIAL ROLE IN RESPECTING THE SPATIAL DISTRIBUTION
CC CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE
CC CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE
CC CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).
CC CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC CC at the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; L27585; AAA20998.1; -
CC CC EMBL; U30711; AAC59742.1; -
CC CC EMBL; Z35669; CAA84738.1; -
CC CC EMBL; AF124382; AAD47913.1; -
CC CC EMBL; U51351; AAB38575.1; -
CC CC EMBL; U51370; AAB38593.1; -
CC CC HSP; Q62226; IVH1.
CC CC MEROPS; C46.002; -.
CC CC ZFIN; ZDB-GENE-980526-166; shh.
CC CC InterPro: IPR000320; HH_signal.
CC CC InterPro: IPR001767; Hedgehog_hhnt.
CC CC InterPro: IPR003586; Hedgehog_hhntc.
CC CC InterPro: IPR003587; Hedgehog_hhntn.
CC CC InterPro: IPR002203; Intein.
CC CC InterPro: IPR001657; SonicHH.
CC CC Pfam; PF01079; Hhnt; 1.
CC CC Pfam; PF01085; HH_signal; 1.
CC CC PRINTS; PR00632; SONICHHOG.
CC CC PRODOM; PD003042; HH_signal; 1.
CC CC SMART; SM00305; Hhntc; 1.
CC CC SMART; SM00306; Hhntn; 1.
CC CC PROSITE; PS50817; INTEIN_N_TER; 1.
CC CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;
CC CC signal; Lipoprotein; Palmitate.
CC CC KW
CC CC SIGNAL; 1
CC CC CHAIN; 23
CC CC CHAIN; 24
CC CC CHAIN; 197
CC CC CHAIN; 198
CC CC CHAIN; 418
CC CC SITE; 197
CC CC SITE; 198
CC CC SITE; 243
CC CC SITE; 243
CC CC SITE; 267
CC CC SITE; 267
CC CC INVOLVED IN AUTO-CLEAVAGE (BY
CC CC SIMILARITY).
CC CC INVOLVED IN AUTO-CLEAVAGE (BY

```

FT ACT\_SITE 270 270 SIMILARITY).  
 FT BINDING 197 197 ESSENTIAL FOR AUTO-CLEAVAGE (BY  
 FT LIPID 24 24 SIMILARITY).  
 FT SEQUENCE 418 AA; 46402 MW; CF000AFEPD25795 CRC64;  
 Query Match 61.0%; Score 1506; DB 1; Length 418;  
 Best Local Similarity 64.5%; Pred. No. 6,7e-104;  
 Matches 300; Conservative 47; Mismatches 68; Indels 50; Gaps 5;

1 MLRLARCLLVVSSLVSLGSLGCGPGGKRRHPRKLTPLAKQFIPNNAEKTGLASG 60  
 1 MRLTRVLVSLVSLVSLGSLGCGPGGKRRHPRKLTPLAKQFIPNNAEKTGLASG 60  
 61 RYEGKISNRSERFELTPNYPNPDIIKDEENTGADRLTORCKDKLNALISVMNOMPGV 120  
 61 RYEGKISNRSERFELTPNYPNPDIIKDEENTGADRLTORCKDKLNALISVMNOMPGV 120  
 61 RYEGKISNRSERFELTPNYPNPDIIKDEENTGADRLTORCKDKLNALISVMNOMPGV 120  
 121 KLRVTEGDEGDHSHSESLAYEGRAVDITTSRDRSKYGLMARLAVEAGDFWYVESKAH 180  
 121 KLRVTEGDEGDHSHSESLAYEGRAVDITTSRDRSKYGLMARLAVEAGDFWYVESKAH 180  
 181 IHCSVKAKNSAASGCGFPGSSATVHLEGGTKLVKDLSPDRVLAADOGRLYSDPLT 240  
 181 IHCSVKAKNSAASGCGFPGSSATVHLEGGTKLVKDLSPDRVLAADOGRLYSDPLT 240  
 241 FLDRDDGAKKVFYVIERPERERLLTNAHLFLVAPHNDSATGEPEASSGSGSPSGALG 300  
 241 FLDRDDGAKKVFYVIERPERERLLTNAHLFLVAPHNDSATGEPEASSGSGSPSGALG 300  
 241 FLDRDDGAKKVFYVIERPERERLLTNAHLFLVAPHNDSATGEPEASSGSGSPSGALG 300  
 301 PRALFASRVPRGQRYVVAERDGRRLIPAASVTLSEEAAGAPILTAOCTLLINRVL 360  
 301 PRALFASRVPRGQRYVVAERDGRRLIPAASVTLSEEAAGAPILTAOCTLLINRVL 360  
 287 --AAVASSVRAGQYVYV---DSSGQLKSYIVORYIEEORGSAFAPYTAGTIVYDRIL 340  
 287 --AAVASSVRAGQYVYV---DSSGQLKSYIVORYIEEORGSAFAPYTAGTIVYDRIL 340  
 361 ASCAAVIEESMARAPRFLAHLAALAPARTDRGCGDGGGDRGGGGRVALTARGA 420  
 361 ASCAAVIEESMARAPRFLAHLAALAPARTDRGCGDGGGDRGGGGRVALTARGA 420  
 341 ASCAAVIEESMARAPRFLAHLAALAPARTDRGCGDGGGDRGGGGRVALTARGA 420  
 341 ASCAAVIEESMARAPRFLAHLAALAPARTDRGCGDGGGDRGGGGRVALTARGA 420  
 421 ADABGAGAT--AGIHMSQLLYOIGTWLDSALHPLGMAVKS 462  
 421 ADABGAGAT--AGIHMSQLLYOIGTWLDSALHPLGMAVKS 462  
 376 --SSRSNATLQOEGVHWYSRLLYQMGTWLDSNMLHPLGMSVNS 418  
 376 --SSRSNATLQOEGVHWYSRLLYQMGTWLDSNMLHPLGMSVNS 418

RESULT 7  
 SHH\_XENLA STANDARD; PRT; 444 AA.  
 AC 092000: 091894;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Sonic hedgehog protein precursor (X-SHH) (VHH-1).  
 GN SHH.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 OC Xenopodidae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN 1;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Intestine;  
 RX MEDLINE=95357169; PubMed=7630736;  
 RA Stoolow M.A., Shi Y.-B.;  
 RT "Xenopus sonic hedgehog as a potential morphogen during embryogenesis  
 and thyroid hormone-dependent metamorphosis";  
 RL Nucleic Acids Res. 23:2555-2562(1995).  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=95401852; PubMed=7671800;  
 RA Ekker S.C., McGrew L.L., Lai C.-J., Lee J.J., von Kessler D.P.,  
 Moon R.T., Beachy P.A.;  
 RT "Distinct expression and shared activities of members of the hedgehog

RT gene family of Xenopus laevis.";  
 RL Development 121:2337-2347(1995).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Notochord;  
 RX MEDLINE=96028338; PubMed=7551564;  
 RA Ruiz i Altaba A., Jessell T.M., Roelink H.;  
 RT "Restrictions to floor plate induction by hedgehog and winged-helix  
 genes in the neural tube of frog embryos.";  
 RL Mol. Cell. Neurosci. 6:106-121(1995).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT AND METAMORPHOSIS. INVOLVED  
 CC IN LIMB FORMATION, PATTERNING OF THE CENTRAL NERVOUS SYSTEM AND  
 CC VENTRAL SOMITE DIFFERENTIATION. INDUCES ECTOPIC CEMENT GLAND  
 CC FORMATION IN EMBRYOS. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH  
 CC TRANSCRIPTION IN ASSOCIATION WITH SMOOTHENED (SMO), TO ACTIVATE THE  
 CC REPRESSORS THE CONSTITUTIVE SIGNALING ACTIVITY OF SMO (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN NOTOCHORD AND NEURAL  
 CC FLOOR PLATE DURING EMBRYOGENESIS. IN TADPOLE, HIGH EXPRESSION IS  
 CC OBSERVED IN PANCREAS/STOMACH, MODERATE EXPRESSION IN TAIL, AND LOW  
 CC EXPRESSION IN INTESTINE, BRAIN, AND HIND LIMB.  
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT THE NEURULA (STAGES 16-17).  
 CC FIRST PEAK OF EXPRESSION AROUND TADPOLE HATCHING (STAGES 33-40).  
 CC HIGH EXPRESSION OBSERVED IN INTESTINE AT THE CLIMAX OF  
 CC MORPHOGENESIS (STAGES 60-62) WHEN INTESTINE EPITHELIAL UNDERGOES  
 CC MORPHOGENESIS.  
 CC -1- INDUCTION: BY THYROID HORMONE.  
 CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L39213; AAC4227.1; -;  
 CC EMBL: U26314; AAA85162.1; -;  
 CC EMBL: L35248; AAA4981.1; -;  
 CC HSSP: Q62226; IVH.  
 CC MEROPS: C46.002; -;  
 CC InterPro: IPR000320; HH\_signal.  
 CC InterPro: IPR001767; Hedgehog\_hlnt.  
 CC InterPro: IPR003586; Hedgehog\_hlntc.  
 CC InterPro: IPR003587; Hedgehog\_hlntc.  
 CC InterPro: IPR002203; Inteln.  
 CC InterPro: IPR001657; SonicHH.  
 CC Pfam: PF01079; Hntc.1.  
 CC Pfam: PF01085; Hntc.1.  
 CC PRINTS: PR00632; SONICHHOG.  
 CC PRODOM: PD003042; HH\_signal; 1.  
 CC SMART: SM00305; Hntc.1.  
 CC SMART: SM00306; Hntcn.1.  
 CC PROSITE: PS50817; INTEIN\_N\_TER; 1.  
 CC Developmental protein. Autocatalytic cleavage; Hydrolase; Protease;  
 CC Signal; Lipoprotein; Palmitate; Repeat.  
 FT SIGNAL 1 24 POTENTIAL.







CC 363 CVAIEBHSMAHRAFAFRLAHLAALPARTDRCGDSGGGCGGVALTAPGAAD 422  
CC 344 CFALVNEQDLAQAKAFRLRLYHSL-----GGPGVQGD----- 376  
CC 423 ABGAGTAGIHWYSGLLIYIGTGLDSEALHPG 456  
DB 377 -----GVHWYSGLLYRIGRLPLPPDSFHPG 402

RESULT 10  
ID IH\_HUMAN STANDARD: PRT: 411 AA.  
AC Q14623; Q043322;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Indian hedgehog protein precursor (IHH) (HMG-2).  
GN IHH.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Tate G., Kishimoto K., Mitsuya T.;  
RT "Expression of Sonic hedgehog and its receptor Patched/Smoothed in  
RL human cancer cell lines and embryonic organs.";  
RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).  
RN [2]  
RP SEQUENCE OF 100-411 FROM N.A.  
RC TISSUE=Fetal lung;  
RX MEDLINE=96070431; PubMed=7590746;  
RA Marigo V., Roberts D.J., Lee S.M.K., Tsukurov O., Levi T.,  
RA Gertler J.M., Epstein D.J., Gilbert J., Copeland N.G., Seidman C.E.,  
RA Jenkins N.A., Seidman J.G., McMahon A.P., Tabin C.C.  
RT Cloning, expression, and chromosomal location of SHH and IHH: two  
RT human homologues of the Drosophila segment polarity gene hedgehog.";  
RL Genomics 28:44-51(1995).  
RN [3]  
RP SEQUENCE OF 124-172 FROM N.A.  
RX MEDLINE=95236997; PubMed=7720571;  
RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K.,  
RA Zhao R., Seidman M.F., Fallon J.F., Beachy P.A.;  
RT "Products, genetic linkage and limb patterning activity of a murine  
RT hedgehog gene.";  
RL Development 120:3339-3353(1994).  
CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
CC PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC)  
CC RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHENED (SMO), TO  
CC ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN  
CC ENDOCHONDRAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH  
CC AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION  
CC OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP) (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
CC TO MEDIATE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN EMBRYONIC LUNG, AND IN ADULT  
CC KIDNEY AND LIVER.  
CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY  
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: AB018076; BAA33523.2; -  
CC EMBL: AB010092; BAA33523.2; JOINED.  
CC EMBL: AB018075; BAA33523.2; JOINED.  
CC EMBL: L38517; AAA62178.1; -  
CC HSSP: Q62226; 1YHH.  
CC MEROPS: C46.003; -  
CC GeneW: HGNC:5956; IHH.  
CC MIM: 600726; -  
CC InterPro: IPR000320; HH\_signal.  
CC InterPro: IPR001767; Hedgehog\_hlnt.  
CC InterPro: IPR003586; Hedgehog\_hlntc.  
CC InterPro: IPR003587; Hedgehog\_hlntc.  
CC InterPro: IPR002203; Intein.  
CC Pfam: PF01079; Hlnt; 1.  
CC Pfam: PF01085; HH\_signal; 1.  
CC SMART: SM00305; Hlntc; 1.  
CC SMART: SM00306; Hlntc; 1.  
CC PROSITE: PS50817; INTEIN\_N\_TER; 1.  
CC Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
CC Signal; Lipoprotein; Palmitate.  
CC SIGNAL 1 27  
CC CHAIN 28 411  
CC CHAIN 28 202  
CC CHAIN 203 411  
CC SITE 202 203  
CC SITE 248 248  
CC SITE 272 272  
CC ACT\_SITE 275 275  
CC BINDING 202 202  
CC LIPID 28 28  
CC CONFLICT 100 100  
CC CONFLICT 246 246  
CC CONFLICT 309 309  
CC SEQUENCE 411 AA; 45188 MW; 14C904A36C344DEA CRC64;  
Query Match 51.3%; Score 1266; DB 1; Length 411;  
Best local Similarity 57.5%; Pred. No. 3.4e-86;  
Matches 266; Conservative 45; Mismatches 86; Indels 66; Gaps 10;  
CC 7 CLLLVSSILVCSGLACGPGGFG-KRRHPKKLTPLAYKQFLPNVAEKTGASGRYEGK 65  
CC 13 CLVLL--LVVPAAGCGGPGRVGSRPRKLVPLAVKQFSPVPEKTLGASGRYEGK 70  
CC 66 ISNSRPFELTPNYPDIIIFKDEENTGADRMTQCKRLNALISVNMQGVLYRLT 125  
CC 71 IARSSRPFELTPNYPDIIIFKDEENTGADRMTQCKRLNALISVNMQGVLYRLT 130  
CC 126 EGDDEGHSSEELAHTEGRAVDITTSDDRSKGMGLARLAVEGPMWYVESAHIHCSV 185  
CC 131 EGDDEGHSSEELAHTEGRAVDITTSDDRSKGMGLARLAVEGPMWYVESAHIHCSV 190  
CC 246 DGAKKFFVYETEPREBLITTAHLIFVAPNHDSTGPEASSGGPSGGLAPRALF 305  
CC 251 PHLRAFOVYETEDPPRRRLATPAHLLETA--DNHT-BPAARF-----RAIF 294  
CC 306 ASVRRAQGVYVAERDGRRLPAVHSTLSEEAAGAPLTAGGTILINVLASGVA 365  
CC 295 ASHYVQGVYVA-----GABGLDPAARYAAS-THVALGATAPLTAKGTILVEYVASCF 349  
CC 366 VIEBHSMAHRAFAFRLAHLAALPARTDRCGDSGGGCGGVALTAPGAADAPG 425

DB 350 AVADHILALAFWPLRLEHSL-----ANGSWTPG 378

QY 426 AGATAGIHWYSQLYQITWLLDSEALPHLCMAVKSXSRGAG 468

DB 379 -----EGVHWYPOLLYXRGLRLLLEGSRPLGMS-----GAG 410

RESULT 11

INH\_MOUSE STANDARD: PRT: 411 AA.

AC P97812: 061724: 38, Created

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Indian hedgehog protein precursor (INH) (HMG-2).

GN INH.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID:10090;

RN [1]

RP SEQUENCE FROM N.A., AND AUTOPROTEOLYTIC CLEAVAGE.

RC STRAIN-CD-1: TISSUE-Kidney.

RX MEDLINE-97236802; PubMed-9079674;

RA Valentini R.P., Brookhiser W.T., Park J., Yang T., Briggs J., Dressler G., Holzman L.B.;

RT "Post-translational processing and renal expression of mouse Indian hedgehog.";

RL J. Biol. Chem. 272:8466-8473(1997).

RN [2]

RP SEQUENCE OF 76-411 FROM N.A.

RC STRAIN-C57BL/6J; TISSUE-Embryo;

RX MEDLINE-9404334; PubMed-7916661;

RA Echeleard Y., Epstein D.J., St Jacques B., Shen L., Mohler J., McMahon J.A., McMahon A.P.;

RT " Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity.";

RL Cell 75:1417-1430(1993).

RN [3]

RP REVISIONS.

RC STRAIN-C57BL/6J;

RA St Jacques B.;

RT Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE OF 124-172 FROM N.A.

RX MEDLINE-95236997; PubMed-7720571;

RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Simandl B.K., Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;

RT "Products, genetic linkage and limb patterning activity of a murine hedgehog gene.";

RL Development 120:3339-3353(1994).

CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF PATTERNING EVENTS DURING DEVELOPMENT. BINDS TO THE PATCHED (PTC) RECEPTOR, WHICH FUNCTIONS IN ASSOCIATION WITH SMOOTHED (SMO), TO ACTIVATE THE TRANSCRIPTION OF TARGET GENES. IMPLICATED IN ENDOCHONDRAL OSSIFICATION: MAY REGULATE THE BALANCE BETWEEN GROWTH AND OSSIFICATION OF THE DEVELOPING BONES. INDUCES THE EXPRESSION OF PARATHYROID HORMONE-RELATED PROTEIN (PTHrP).

CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).

CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, DETECTED IN THE DEVELOPING GUT, AND UROGENITAL SINUS. IN THE ADULT KIDNEY, FOUND IN PROXIMAL CONVOLUTED AND PROXIMAL STRAIGHT TUBULE.

CC -1- DEVELOPMENTAL STAGE: DETECTED AT 10 DAYS POST COITUM (DPC) IN DEVELOPING GUT. AT 14.5 DAYS DPC IN THE CARTILAGE PRIMORDIUM AND IN THE DEVELOPING UROGENITAL SINUS. EXPRESSION INCREASES WITH GESTATIONAL AGE IN KIDNEY AND DUODENUM, BECOMING MAXIMAL IN ADULTHOOD.

CC -1- PTM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN

CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC -----

DR EMBL: U85610; AAB49692.1; ALT\_INIT.

DR EMBL: X76291; CAA53923.1; -.

DR HSSP: Q62226; 1VHH.

DR MEROPS: C46.003; -.

DR MGD: MGI:965533; 1bh.

DR InterPro: IPR000320; HH\_signal.

DR InterPro: IPR001767; Hedgehog\_hnt.

DR InterPro: IPR003586; Hedgehog\_hntc.

DR InterPro: IPR003587; Hedgehog\_hntn.

DR InterPro: IPR002203; Intein.

DR Pfam: PF01079; Hint; 1.

DR Pfam: PF01085; HH\_signal; 1.

DR ProDom: PD003042; HH\_signal; 1.

DR SMART: SM00305; Hntc; 1.

DR SMART: SM00306; Hntn; 1.

DR PROSITE: PS00817; INTEIN\_N\_TER; 1.

DR Developmental protein: Autocatalytic cleavage; Hydrolyase; Protease.

KW Signal; Lipoprotein; Palmitate.

FT SIGNAL 1 27

FT CHAIN 28 411

FT CHAIN 28 202

FT CHAIN 203 411

FT SITE 202 203

FT SITE 248 248

FT SITE 272 272

FT ACT\_SITE 275 275

FT BINDING 202 202

FT LIPID 28 28

FT CONFLICT 383 383

FT SEQUENCE 411 AA; 45485 MW; 08BE7AD8507C0D9B CRC64;

Query Match 51.1%; Score 1262; DB 1; Length 411;

Best Local Similarity 56.9%; Pred. No. 6.6e-86;

Matches 264; Conservative 46; Mismatches 86; Indels 68; Gaps 10;

QY 7 CLLVLYSSLLVYSGLAGCGRGFG-KRRHPPKLLPLAYKOFIPNVAEKTGASRGYEGK 65

DB 13 CLFLLLL-LLVPAARGCGPRVYVGRRRPRKLVPLAYKQFSPNVETKLGASRGYEGK 70

QY 66 ISRSSEKREKLTPTVNDITIKDEENTGADRLMTORCKDKNALALISVNMOPGVKRLVT 125

DB 71 IARSSERKELTPTVNDITIKDEENTGADRLMTORCKDKNALALISVNMOPGVKRLVT 130

QY 126 EGMDEDGHHSESLHYGSAVDITTSDDRSKYGLMLRLAYEAGFDWVYYSKAHICSV 185

DB 131 EGMDEDGHHSESLHYGSAVDITTSDDRSKYGLMLRLAYEAGFDWVYYSKAHICSV 190

QY 186 KAENSVAKSGCGPGSAVHLEOGGTRKYVLDLSPGCRVLAADOGRLLYSDFTLFDRD 245

DB 191 KSEHSAAKGCGCPAGAOVRLNGERVALSAVAPGRVLAAMGDGTPTEFDVILFDRE 250

QY 246 DGAKKVFYVITREPRERLLTAHLLFVAPDHNSATGEPASSGSGPSCGALGPALF 305



Db 251 PNLRAFOYIETODPPRRLLATPAILLFTA---DNHT-EPAHF-----RATF 294  
 Qy 306 ASHVRPGORYVVAERGDRLRLPAVHVSUTLSEEAAGAPLTAOCTILINRLVASCYA 365  
 Db 295 ASHVQGOYVYV---SGVPGIOPARVAASV--THVALGSAPLIRHGTILVEEDVVASCA 349  
 Qy 366 VIEHSHANRAFAPEFLAHALL-ALALAPARTDRGSGGGGGGGRGVALTPAGAADAP 424  
 Db 350 AVADHHLAQLAEFWPLRLPFLSLANGSWTPSP-----379  
 Qy 425 GAGATAGIHWSQLYQIGTWLDSALHPILGMAVKSXSXRGAG 468  
 Db 380 -----GVHWPQMLYRLGLRLLESTFHPLGMS-----GAG 410  
 RESULT 12  
 ID IHH\_XENLA STANDARD; PRT; 409 AA.  
 AC 091612;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Indian hedgehog protein precursor (IHH) (Banded hedgehog protein) (X-BHH).  
 GN IHH OR BHH.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 NC NCBL\_TaxID=8355;  
 RX MEDLINE=95401852; Pubmed=7671800;  
 RA Ecker S.C., McGrew L.L., Lal C.-J., Lee J.J., von Kessler D.P.,  
 RA Moon R.T., Beachy P.A.;  
 RT "Distinct expression and shared activities of members of the hedgehog gene family of Xenopus laevis";  
 RL Development 121:2337-2347(1995).  
 CC -1- FUNCTION: SIGNAL INVOLVED IN THE EARLY INDUCTION AND PATTERNING OF ANTERODORSAL ECTODERM, NERVOUS SYSTEM AND SOMITES. INDUCES ECTOPIC CEMENT GLAND FORMATION IN EMBRYOS.  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE CELL SURFACE. HEDGEHOG PROTEIN IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- DEVELOPMENTAL STAGE: DETECTABLE WITHIN THE EARLY GASTRULA, AT STAGE 14 NEURULA, HIGH EXPRESSION IN NEUROECTODERM. EXPRESSED THROUGHOUT THE NEURAL PLATE AND SUBSEQUENTLY IN BOTH THE NERVOUS SYSTEM AND IN THE DERMATOME OF SOMITES.  
 CC -1- INDUCTION: BY ACTIVIN.  
 CC -1- PPM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPHOREOLYSIS ACTIVITY AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: U26404; AAA85165.1; -  
 DR HSSP: 062226; IHH.

DR MEROPS; C46.UFW; -  
 DR InterPro: IPR000320; HH\_signal.  
 DR InterPro: IPR001767; Hedgehog\_hlnc.  
 DR InterPro: IPR003586; Hedgehog\_hlnc.  
 DR InterPro: IPR003587; Hedgehog\_hlnc.  
 DR InterPro: IPR001657; SonicHH.  
 DR Pfam: PF01079; Hlnc; 1.  
 DR Pfam: PF01085; HH\_signal; 1.  
 DR PRINTS: PR00632; SONICHHOG.  
 DR Prodom: PD003042; HH\_signal; 1.  
 DR SMART: SM00305; Hlnc; 1.  
 DR SMART: SM00306; Hlnc; 1.  
 KW Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;  
 KW signal; Lipoprotein; Palmitate.  
 FT SIGNAL 1 23  
 FT CHAIN 24 409  
 FT CHAIN 24 197  
 FT CHAIN 198 409  
 FT SITE 197 198  
 FT SITE 267 267  
 FT ACT\_SITE 270 270  
 FT BINDING 197 197  
 FT DOMAIN 32 35  
 FT LIPID 24 24  
 SQ SEQUENCE 409 AA; 45591 MW; 6FB265367FB9627 CRC64;  
 Query Match 50.5%; Score 1248; DB 1; Length 409;  
 Best Local Similarity 56.3%; Pred. No. 7.1e-85;  
 Matches 256; Conservative 53; Mismatches 92; Indels 54; Gaps 7;  
 Qy 4 LARCLLVIVSSLLVSCGL-ACGPGKGRGRRRPPKLTPLATYQFTPNVAETLGSAGRY 62  
 Db 3 LPRVLLICAAALLLGAVNGCCPGKRVGRRRPTLSPSTYQSPNVEKTLGASGRY 62  
 Qy 63 EGRISRNSERFKELTPYNDIIFKDEENTGADRLTORCKRLNALAISVMNQMPGVKL 122  
 Db 63 EGRISRNSERFKELTPYNDIIFKDEENTGADRLTORCKRLNALAISVMNQMPGVKL 122  
 Qy 123 RYEGWDEDHHEESLHYGKRAVDITTSRDSKTKGMLARLAVEAGFDWYVESAHNH 182  
 Db 123 RYEGWDEDHHEESLHYGKRAVDITTSRDSKTKGMLARLAVEAGFDWYVESAHNH 182  
 Qy 183 CSVKAENSVAKGCGEPGSAFVHLBOGGKTLKVDISPGRVLAADQGLVSDFLTFPL 242  
 Db 183 CSVKAENSVAKGCGEPGSAFVHLBOGGKTLKVDISPGRVLAADQGLVSDFLTFPL 242  
 Qy 183 CSVKAENSVAKGCGEPGSAFVHLBOGGKTLKVDISPGRVLAADQGLVSDFLTFPL 242  
 Db 243 DRDDGAKKPYVYIETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSPPSGALGPR 302  
 Qy 243 DRDDGAKKPYVYIETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSPPSGALGPR 302  
 Db 243 DRDDGAKKPYVYIETREPRERLLLTAHLLFVAPHNDSATGEPEASSGSPPSGALGPR 302  
 Qy 303 ALPASRVPGORYVVAERGDRLRLPAVHVSUTLSEEAAGAPLTAOCTILINRLVASCYA 362  
 Db 287 AVPAASVPRQOYI-LVSNVG---LIPAKVRSV-NQTNGAAPLPHOHTLVDDVVS 341  
 Qy 363 CVYVIEHSHANRAFAPEFLAHALLALAPARTDRGSGGGGGGGRGVALTPAGAADAP 422  
 Db 342 CFWLVQKQRLAQLVWPLRLVNL-----GIAGTOP-----373  
 Qy 423 APGAGATAGIHWSQLYQIGTWLDSALHPILGMAVKSXSXRGAG 468  
 Db 374 -----SOOMGIHWSKALYHLGLRLHNGNEHPHIGI 404  
 RESULT 13  
 ID IHH\_BRARE STANDARD; PRT; 412 AA.  
 AC 098862;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Indian hedgehog protein precursor (IHH) (Echidna hedgehog protein).



RP SEQUENCE OF 120-168 FROM N.A.  
 RA MEDLINE:95336997; PubMed-7720571;  
 RA Chang D.T., Lopez A., von Kessler D.P., Chiang C., Smandl B.K.,  
 RA Zhao R., Seidlin M.F., Fallon J.F., Beachy P.A.;  
 RT "Products, genetic linkage and limb patterning activity of a murine  
 RT hedgehog gene.";  
 RL Development 120:3339-3353(1994).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC SPERMATOCTE SURVIVAL DURING DEVELOPMENT. MAY FUNCTION AS A  
 CC PATTERNING EVENTS DURING DEVELOPMENT. ESSENTIAL FOR TESTES  
 CC DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE  
 CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
 CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTES. NOT EXPRESSED IN  
 CC LIMB BUDS.  
 CC -1- PIM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPHOREOLYSIS ACTIVITY  
 CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
 CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
 CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
 CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
 CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
 CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
 CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
 CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
 CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch))  
 CC -----  
 CC EMBL: X76292; CAA53924.1; -;  
 CC DR HSSP: Q62226; LVNH.  
 CC DR MEROPS: C46.004; -;  
 CC DR MGD: MGI:94891; Dhh.  
 CC DR InterPro: IPR000320; HH\_signal.  
 CC DR InterPro: IPR001767; Hedgehog\_hhnc.  
 CC DR InterPro: IPR003586; Hedgehog\_hhnc.  
 CC DR InterPro: IPR002203; Intein.  
 CC DR InterPro: IPR001657; Sonichh.  
 CC DR Pfam: PF01079; Hhnt; 1.  
 CC DR Pfam: PF01085; Hh\_signal; 1.  
 CC DR PRINTS: PRO0632; SONICHHOG.  
 CC DR PRODOM: PD003042; HH\_signal; 1.  
 CC DR SMART: SM00305; Hhnc; 1.  
 CC DR SMART: SM00306; Hhnt; 1.  
 CC DR PROSITE: PS50817; INTEIN\_NTER; 1.  
 CC DR Developmental protein: Autocatalytic cleavage; Hydrolase; Protease;  
 CC Signal; Lipoprotein; Palmitate.  
 CC KW  
 CC FT SIGNAL 1 22  
 CC FT CHAIN 23 396  
 CC FT CHAIN 23 198  
 CC FT CHAIN 199 396  
 CC FT SITE 198 199  
 CC FT SITE 244 244  
 CC FT SITE 244 244  
 CC FT SITE 268 268  
 CC FT ACT\_SITE 271 271  
 CC FT BINDING 198 198  
 CC FT LIPID 23 23  
 CC FT SEQUENCE 396 AA; 43542 MW; AFEB051BE950FD8 CRC64;  
 Query Match 47.9%; Score 1183; DB 1; Length 396;  
 Best Local Similarity 54.5%; Pred. No. 4, 3e-80;  
 Matches 244; Conservative 55; Mismatches 87; Indels 62; Gaps 10;

QY 3 LIAICLLLVVSSLLVCSGLACGPGC-FGKRRHPRK-LTPLAYKQFIPNVAEKTGASG 60  
 DB 7 LIPICLL-----ALLALSQSCGPGRGVGRRRYRKQLVPLLYKQFVSMBERTLGASG 61  
 QY 61 RYEGKSRNSRSEFKELTPYNPDIIFKDEENCGADRLMTORCKDKLINALAISYMNQPGV 120  
 DB 62 PAGRVTNRSERFRDLVPYNPDIIFKDEENSGADRLMERCKERYNALAIIVMNMPCV 121  
 QY 121 KLRVTEGMEDEHSHSESHYEGRAVDITSPDRSKYGMRLRAVEAGFDWYYESKAH 180  
 DB 122 RLRYTEGMEDEHSHHADSDHYEGRALDITTSDRDKKIGLLRLAVEAGFDWYYESRHH 181  
 QY 181 HCSYVAENSVAASGCGCPGSAATVHLDEGGRKLYKVDLSPGDRVLAADOGRLYSDFLT 240  
 DB 182 IHVSVAADSLAVRAGGCPGSAATVRLRSGERKGLRELRGDMVLAADAAGVVPVLL 241  
 QY 241 FLDRDGAARVYVETREPRERLITLAHLFVAPHNDSATGEPAASSGCP-PSGCL 299  
 DB 242 FLDRDGAARVYVETREPRERLITLAHLFVAPHNDSATGEPAASSGCP-PSGCL 299  
 QY 300 GPRALFASVRPGQRYVVAERDGRRLPAVHSVTLSEEAAGVAPLTAGTILINRY 359  
 DB 287 AP--VFARLRAGDSVLA---PGDALQPARVARYA-REAVGVFAPLTAGTILINRY 339  
 QY 360 LASCYAVIEHSVNRAPAFRLAHLAALAPARTDRGDSGCGGCGGCVALTAPG 419  
 DB 340 LASCYAVIEHSVNRAPAFRLAHLAALAPARTDRGDSGCGGCGGCVALTAPG 419  
 QY 420 AADPAGATAGIHWYSQLYQIGTWL 447  
 DB 373 GAVQP-----TGMHWYSRLYLAEELM 395  
 RESULT 15  
 DHH\_HUMAN STANDARD; PRT; 396 AA.  
 AC 043323; Q15794;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Desert hedgehog protein precursor (DHH) (HhG-3).  
 GN DHH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Tate G., Kishimoto K., Mitsuura T.;  
 RT "Expression of Sonic hedgehog and its receptor patched/smoothed in  
 RT human cancer cell lines and embryonic organs.";  
 RL J. Biochem. Mol. Biol. Biophys. 4:27-34(2000).  
 RN [2]  
 RN SEQUENCE OF 85-178 FROM N.A.  
 RC TISSUE=Kidney;  
 RA Drummond I.A.;  
 RT "Human desert hedgehog";  
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RN DISEASE;  
 RX PubMed-11017805;  
 RA Umebara F., Tate G., Itoh K., Yamaguchi N., Douchl T., Mitsuura T.,  
 RA Osame M.;  
 RT "A novel mutation of desert hedgehog in a patient with 46,XY partial  
 RT gonadal dysgenesis is accompanied by multifascicular neuropathy.";  
 RL Am. J. Hum. Genet. 67:1302-1305(2000).  
 CC -1- FUNCTION: INTERCELLULAR SIGNAL ESSENTIAL FOR A VARIETY OF  
 CC PATTERNING EVENTS DURING DEVELOPMENT. MAY FUNCTION AS A  
 CC SPERMATOCTE SURVIVAL FACTOR IN THE TESTES. ESSENTIAL FOR TESTES  
 CC DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: THE C-TERMINAL PEPTIDE DIFFUSES FROM THE  
 CC CELL, WHILE THE N-TERMINAL PEPTIDE REMAINS ASSOCIATED WITH THE

CC CELL SURFACE. IS ALSO SECRETED IN EITHER CLEAVED OR UNCLEAVED FORM  
CC TO MEDIANE SIGNALING TO OTHER CELLS (BY SIMILARITY).  
CC  
CC -1- PM: THE C-TERMINAL DOMAIN DISPLAYS AN AUTOPROTEOLYTIC ACTIVITY  
CC AND A CHOLESTEROL TRANSFERASE ACTIVITY. BOTH ACTIVITIES RESULT IN  
CC THE CLEAVAGE OF THE FULL-LENGTH PROTEIN AND COVALENT ATTACHMENT OF  
CC A CHOLESTEROL MOIETY TO THE C-TERMINAL OF THE NEWLY GENERATED N-  
CC TERMINAL FRAGMENT (N-PRODUCT). THIS COVALENT MODIFICATION APPEARS  
CC TO PLAY AN ESSENTIAL ROLE IN RESTRICTING THE SPATIAL DISTRIBUTION  
CC OF THE PROTEIN ACTIVITY TO THE CELL SURFACE. THE N-PRODUCT IS THE  
CC ACTIVE SPECIES IN BOTH LOCAL AND LONG-RANGE SIGNALING, WHEREAS THE  
CC C-PRODUCT HAS NO SIGNALING ACTIVITY (BY SIMILARITY).  
CC  
CC -1- DISEASE: Defects in DHH may be the cause of partial gonadal  
CC dysgenesis (PGD) with multifactorial neuropathy. PGD is  
CC characterized by the presence of a testis on one side and a streak  
CC or an absent gonad at the other, persistence of Mullerian duct  
CC structures, and a variable degree of genital ambiguity.  
CC  
CC -1- SIMILARITY: BELONGS TO THE HEDGEHOG FAMILY.  
CC -----  
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL:	AB010994:	BAA24866.1:	-	JOINED.
DR	EMBL:	AB010581:	BAA24866.1:	JOINED.	
DR	EMBL:	AB010993:	BAA24866.1:	JOINED.	
DR	HSSP:	U59748:	AA03398.1:	-	
DR	HSSP:	O62226:	1VHH.		
DR	MEROPS:	C46_004:	-		
DR	Genew:	HGNC:2865:	DHH.		
DR	MTM:	605423:	-		
DR	MTM:	607080:	-		
DR	InterPro:	IPR000320:	HH_signal.		
DR	InterPro:	IPR001767:	Hedgehog_hnt.		
DR	InterPro:	IPR003586:	Hedgehog_hntc.		
DR	InterPro:	IPR003587:	Hedgehog_hntnc.		
DR	InterPro:	IPR001657:	SonicHH.		
DR	Pfam:	PF01079:	Hntc.1.		
DR	Pfam:	PF01085:	HH_signal.1.		
DR	PRINTS:	PR00632:	SONICHHOG.		
DR	Prodom:	PD003042:	HH_signal.1.		
DR	SMART:	SM00305:	Hntc.1.		
DR	SMART:	SM00306:	Hntnc.1.		
KW	Developmental protein; Autocatalytic cleavage; Hydrolase; Protease;				
KW	Signal; Lipoprotein; Palmitate.				
FT	SIGNAL	1	22		POTENTIAL.
FT	CHAIN	23	396		DESERT HEDGEHOG PROTEIN.
FT	CHAIN	23	198		DESERT HEDGEHOG PROTEIN N-PRODUCT.
FT	CHAIN	199	396		DESERT HEDGEHOG PROTEIN C-PRODUCT.
FT	STATE	198	199		CLEAVAGE (AUTO-) (BY SIMILARITY).
FT	STATE	244	244		INVOLVED IN CHOLESTEROL TRANSFER (BY SIMILARITY).
FT	STATE	268	268		INVOLVED IN AUTO-CLEAVAGE (BY SIMILARITY).
FT	ACT_SITE	271	271		ESSENTIAL FOR AUTO-CLEAVAGE (BY SIMILARITY).
FT	BINDING	198	198		CHOLESTEROL (BY SIMILARITY).
FT	LIPID	23	23		PALMITATE (BY SIMILARITY).
FT	CONFLICT	177	177		E -> G (1N REF.2).
SQ	SEQUENCE	396 AA:	43577 MR:		FCE4FB21972C3AD5 CRC64:

Query Match	47.7%	Score 1177	DB 1	Length 356
Best Local Similarity	54.4%	Pred. No. 1.2e-79		
Matches 245	Conservative 54	Mismatches 93	Indels 58	Gaps 10

```

0Y      1 MLLARCLLVYSSLLVCSGLACGSPRG-FGKRHRHK-LTPLYAQFIPNVAETLGA 58
      | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db      1 MALLTNLLPLCC-LALLALPAQSGSPRGPGVGRRRYARKQVPLLYXQFPGVPERLGA 59
0Y      59 SGREYKISIRNSERFELTPYNYNDILFKDEENTGADRLMTQRCOKDLNALATISVMNQMP 118

```

Db	60	SGPAGRGARARSEFPRDLVPIYNYNDIIIFKDENSGABRLMTERCKERNMALATAVNMMP	11.9%
QY	119	GVKLTETGEMDEDGHHSSESLHYEGRAVDITTSRDRSKYGMALARLAVAGFDWVYVESK	17.8%
Db	120	GVRLRVETGEMDEDGHHSADSLHYEGRALDITTSRDRSKYGMALARLAVAGFDWVYVESR	17.9%
QY	179	AHICSVYAEVSVAKASGCGEPFSATVHLDEGGRVLYKDDSPGDRVLAADOGRLVYSDF	23.8%
Db	180	NHNVSVYKADNSLAVRAGCGEPFNATVRLMNGEKKGLRELHGRDWTVLAADSGHVVTPV	23.8%
QY	239	LTFELDRDDGAKVVFYIETREPRERLLLTAAHLFVAPRHNDSATGEPEASSGSP-PSGG	29.7%
Db	240	LLFLFDRDLORRASEVAVETEMPRKLLLTWNHLVFAA-----RGAPAPAC	28.4%
QY	298	ALGRPLFASVRVREGORYVVAEDSGDRLLRPAVHSYLTSEAAAGVAPRLTAGDTLLN	35.5%
Db	285	DFAP--VFARKLKRGDSVLA-----PGGDALRPARYARA--RELVAGFAPRLTAGDTLLN	33.7%
QY	358	RVLASCAVIEHSHMAHRAFAFRLIHALLAALAPARTDGGDSGGDRGGGGGRVALTA	41.7%
Db	338	DVLASCAVYIEHSHMAHRAFAFRLIHLNIGALL-----	37.0%
QY	418	PGAADAPGAGATAGIHHYSOLLXOIGTWLL	44.7%
Db	371	PGGAVQP-----TGHHYTSRLYLRLEELL	39.5%

```
Search completed: February 20, 2003, 10:12:40
Job time : 15.5662 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:06:28 ; Search time 30.5535 Seconds  
(without alignments)  
3203.314 Million cell updates/sec

Title: US-09-827-110A-15

Perfect score: 2469  
Sequence: 1 MLLARCLLVSLVCS.....GMAYKSKRGAGGAREGA 475

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_21.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1544.5	62.6	434	13	057404 pleurodeles
2	1454	58.9	414	13	09W709
3	1308	53.0	406	13	057567
4	1094.5	44.3	415	5	090526
5	1084.5	43.9	415	5	017499
6	1049.5	42.5	442	13	073803
7	1041.5	42.2	447	5	09G8A8
8	916	37.1	410	5	061676
9	850	34.4	161	11	09R179
10	733	29.7	150	11	09IGV7
11	731	29.6	139	6	09X516
12	709	28.7	138	11	09W6C1
13	702	28.4	177	11	09WV29
14	619	25.1	185	5	096699
15	531	21.5	99	13	080GNA
16	523	21.2	119	13	042128

17	511	20.7	129	11	09WUP6	09WUP6 rattus norv
18	476.5	19.3	137	13	042234	042234 coturnix co
19	392	15.9	80	13	042441	042441 oryzias lat
20	392	15.9	87	5	09TX30	09TX30 anopheles g
21	339	13.7	63	13	091814	091814 rana catesb
22	332	13.4	64	13	09PRF5	09PRF5 oryzias lat
23	298	12.1	60	5	0952D8	0952D8 hydra atten
24	289	11.7	56	13	0902R0	0902R0 oryzias lat
25	237	9.6	49	5	09TX33	09TX33 hindu medl
26	229	9.3	45	5	09TX31	09TX31 tribolium c
27	188	7.6	54	13	042233	042233 coturnix co
28	187.5	7.6	125	11	09ESH3	09ESH3 rattus norv
29	175.5	7.1	48	5	09TX32	09TX32 strongyloce
30	163.5	6.6	557	5	094410	094410 caenorhabd1
31	163.5	6.6	557	5	094129	094129 caenorhabd1
32	159.5	6.5	615	5	P91573	P91573 caenorhabd1
33	149	6.0	1207	5	021535	021535 caenorhabd1
34	138.5	5.6	790	5	022872	022872 caenorhabd1
35	138.5	5.6	1203	5	021835	021835 caenorhabd1
36	135.5	5.5	485	5	094128	094128 caenorhabd1
37	134.5	5.4	481	5	045992	045992 caenorhabd1
38	128	5.2	550	5	045273	045273 caenorhabd1
39	126.5	5.1	1021	5	09XUV2	09XUV2 caenorhabd1
40	125	5.1	550	5	094130	094130 caenorhabd1
41	123.5	5.0	2639	5	076786	076786 antherea p
42	121	4.9	868	5	09XV14	09XV14 caenorhabd1
43	119	4.8	318	16	086703	086703 streptomyce
44	119	4.8	2655	5	0964F4	0964F4 antherea y
45	118	4.8	1832	3	08TGH8	08TGH8 podospora a

## ALIGNMENTS

### RESULT 1

ID	057404	PRELIMINARY;	PRT;	434 AA.
AC	057404;			
DT	01-JUN-1998 (TREMBLrel. 06, Created)			
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Sonic hedgehog-related protein.			
GN	tw-shh.			
OS	Pleurodeles waltlil (Iberian ribbed newt).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;			
OC	Pleurodeles.			
OX	NCBI_TaxID=8319;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Caubit X., Nicolas S., Le Parco Y.;			
RT	"Pleurodeles sonic hedgehog";			
RL	Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF003532; AAB94412.1; -.			
DR	HSSP; Q62226; 1VH.			
DR	MEROPS; C46.002; -.			
DR	InterPro: IPR001767; Hedgehog_hntc.			
DR	InterPro: IPR003586; Hedgehog_hntc.			
DR	InterPro: IPR003587; Hedgehog_hntc.			
DR	InterPro: IPR003320; Hh_signal.			
DR	InterPro: IPR002203; Intein.			
DR	InterPro: IPR001657; SonicHH.			
DR	Pfam: PF01085; Hh_signal; 1.			
DR	Pfam: PF01079; Hntc; 1.			
DR	PRINTS; PRO0632; SONICHHOG.			
DR	PRODOM; PD003042; Hh_signal; 1.			
DR	SMART; SM00305; Hntc; 1.			
DR	SMART; SM00306; Hntc; 1.			
DR	SEQUENCE 434 AA; 48421 MW; A9495E367151AE74 CRC64;			

Query Match 62.6%; Score 1544.5; DB 13; Length 434;  
Best Local Similarity 67.7%; Pred. No. 8.1e-108;  
Matches 313; Conservative 41; Mismatches 77; Indels 31; Gaps 6;

[illegible]

RESULT 2			
Q9W709	PRELIMINARY;	PRT;	414 AA.
ID Q9W709			
AC Q9W709;			
DT 01-NOV-1999 (TREMBLrel. 12, Created)			
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE Sonic hedgehog.			
GN Shh.			
OS Paracichthys olivaceus (Flounder).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;			
OC Pleuronectoidae; Paracichthyidae; Paracichthys.			
OX NCBI_TaxID=6255;			
RN [1]			
SEQUENCE FROM N.A.			
RA MEDLINE-99238226; PubMed-10223710;			
RA Suzuki T., Ichiro O., Kurokawa T.;			
RT "Retinoic acid given at late embryonic stage depresses sonic hedgehog			
RT and Hoxd-4 expression in the pharyngeal area and induces skeletal			
RT malformation in flounder (Paracichthys olivaceus) embryos.";			
RL Dev. Growth Differ. 4:143-152(1999).			
DR EMBL; AB029748; BAA82360.1; -			
DR HSSP; Q62226; 1VHH. -			
DR MEROPS; C46.001; -			
DR InterPro; IPR001767; Hedgehog_hntc.			
DR InterPro; IPR003586; Hedgehog_hntc.			
DR InterPro; IPR003587; Hedgehog_hntc.			
DR InterPro; IPR003200; HH_signal.			
DR InterPro; IPR002203; Intein.			
DR InterPro; IPR001657; SonicHH.			
DR Pfam; PF01085; HH_signal; 1.			
DR Pfam; PF01079; Hntc; 1.			
DR PRINTS; PR00632; SONICHOG.			
DR ProDom; PD003042; HH_signal; 1.			
DR SMART; SM00305; Hntc; 1.			
DR SMART; SM00306; Hntc; 1.			
SO SEQUENCE 414 AA; 45945 MW; 50607BF3DB7C0DA3 CRC64;			

Query Match	58.9%	Score 1454;	DB 13;	Length 414;
Best Local Similarity	62.4%	Pred. NO. 4.7e-101;		
Matches 289; Conservative	55;	Mismatches 69;	Indels 50;	Gaps 77;

```

OY 1 MLLARCLLLVSSLLVSCSLACPGPRGPKRRHPKPLPLAYKOPTIPVAEKTGLGASG 60
Db 1 MLMLTRIVLAGVITCLSLVSSMGCGPRGRRHPKPLPLAYKOPTIPVAEKTGLGASG 60
OY 61 RYEGKISHNSRFEKELTPVNPNDIIFKDEENTGADRLMTORCKDKLALASVNMOPGV 120
Db 61 RYEEKITRNSRFEKELTPVNPNTDIIIFKDEENTGADRLMTORCKDKLNSIASVNMOPGV 120
OY 121 KLRTIEGMDDEGHHSSESLHTEGAANDITSSDRSRKTYGMLARLAVAGDWTYYESKAH 180
Db 121 KLRTIEGMDDEGHHEESLHTEGAANDITSSDRKSXYGTL SRLAVAGCDWYYESKAH 180
OY 181 IHCYKAKENSVAAKSGCGCPGSAFVHLEOGGCKVYKOLSPGDRYLANDDOGRLLSPDLT 240
Db 181 IHCYKAKENSVAAKSGCGCPGSSVTYLODDGKKRYKALQTDGRYLANDADAGCPYTFDIM 240
OY 241 FLDRDDGAKKVFYIETREPRERLLTAAHLLEVPAPNDSATGCEPAASSGSGPPSGALG 300
Db 241 FIDDDSTTRRLFYIET--DSGOKITTLTAHLLFFGHNSNTERAH-----RG 285
OY 301 PRALEFASVRPGGVYVAERDGRRLTLPAAVHSVLTSEEAAGVAPLTAOGTILIMRYL 360
Db 286 MSAPFAAGVRRGQVYFVPL--DAE-RLQPPYVKRI-YTQHEGSGFAVYTAQGVVYDQVL 340
OY 361 ASCYAIVIEHSMAHRAAPFRLAHALLAALAPARTDRGDSGGGDRGGGGRVALTAPGA 420
Db 341 ASCYAIVIODHELAMALAPVRLAHMVSLLFFSSQ----- 374
OY 421 ADAPGAGATA-GIHWSSOLLYOIGTWLLDSEALHPILGMAYKSS 462
Db 375 ---PQASQKQDGVHWYSKILQTLSTWLLDSHSHPILMAYSPS 414

```

RESULT 3	057567	PRELIMINARY;	PRT;	406 AA.
ID	057567			
AC	057567;			
DT	01-JUN-1998 (TrEMBLrel. 06, Created)			
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Hedgehog segment polarity homolog.			
OS	Notoptthalmus viridescens (Eastern newt) (Triturus viridescens).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae;			
OC	Notoptthalmus.			
OX	NCBI_TaxID=8316;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Stark D.R., Gates P.B., Brookes J.P., Ferretti P.;			
RT	"Hedgehog homologue from Notoptthalmus viridescens.";			
RL	Dev. Dyn. 0:0-0(1998)			
CC	-1 SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE PHOSPHORIBOSYLTRANSFERASE FAMILY.			
CC	EMBL: AF047466; AAC03108.1; -. HSSP: Q62226; 1VHH.			
DR	InterPro: IPR001767; Hedgehog_hint.			
DR	InterPro: IPR003586; Hedgehog_hintC.			
DR	InterPro: IPR003587; Hedgehog_hintN.			
DR	InterPro: IPR000320; HH_signal.			
DR	InterPro: IPR002375; pr/py_tp_transf.			
DR	InterPro: IPR001657; SonicHH.			
DR	Pfam: PF01085; HH_signal; 1.			
DR	Pfam: PF01079; Hint; 1.			
DR	PRINTS: PR00632; SONICHHOG.			
DR	ProDom: PD003042; HH_signal; 1.			
DR	SMART: SM00305; HintC; 1.			
DR	SMART: SM00306; HintN; 1.			
DR	PROSITE: PS00103; PUR_PYR_PR_TRANSFER; 1.			

KW Glycoamyltransferase; Transferase  
SQ SEQUENCE 406 AA; 45072 MW; 5842CCAD531AD4D0 CAC64;  
  
Query Match 53.0%; Score 1308; DB 13; Length 406;  
Best Local Similarity 59.2%; Pred. No. 4,2e-90;  
Matches 274; Conservative 41; Mismatches 80; Indels 68; Gaps 9;

```

Query Match          44.3%; Score 1094.5; DB 5; Length 415;
Best Local Similarity 50.2%; Pred. No. 4.7e-74;
Matches 231; Conservative 57; Mismatches 111; Indels 61; Gaps 10;

OY 3 ILARLL-LIVYSSILVCSG--LACGPGSGPKRRHPPKLPFLAYKQFIPIPAEKTGLAS 59
DB 4 VLAARMVTLVALTALGTTHWGSEACPGGTRGRRRHPRKLPFLPVYKQOPASVNTFGAS 63
OY 60 GYEGEKISRNSERFELTPYNPDIIFKDEENTGADRLTQRCXKRLNALAISVNQMPG 119
DB 64 GLFNGRITPDSRRFHTLKNQNFNTDILFDEETGADRFRTQRCXKRLNALAISVNQMPG 123
OY 120 VKLRTYEGWDEGDHHSSESLAYEGRAVDITTSDRDSKXGMLARLAVEAGPDWYYESKA 179
DB 124 VKLRTYEGWDEGDHHSSESLAYEGRAVDITTSDRDSKXGMLARLAVEAGPDWYYESKA 183
OY 180 HHCSCYKANENSAASGCGFPGSATVYHLEOGGTGLVKDLSPEDDRVLADDOGRLLYDPL 239
DB 184 HHCSCYKASDPTAATOGCGFPRESWTRDGRGRIKRRDRVPDKVLSMDSACHPVFSEVL 243
OY 240 TFLDRDGGCKKFEVYIETREPRERLLLTAAHLFLVAPRHDSATGEPEASSGSGPPSGAL 299
DB 244 TFLDRDGRKRPWYITTKTDRNITVTATPASHLYFTESRDL----- 285
OY 300 GPR-ALFASRVPRGQRYVYVAERGD--RRLPAAVHSVTLSEEAAGAYAPLTAOGTILI 356
DB 286 ABRVAKFMSDVAPRGE-FLTPESDGGGFRKY--EIVSYTMEE-KGAVAPLTVHGTVYV 340
OY 357 NRYLASCVAVIEEHSNAHAFAFPRIAAHLALAPARTDRCGDSGGGRCGGCGRVALT 416
DB 341 DMVAWSCAALISSQALAHVFAFPRFLYLYLTSSL----- 374
OY 417 AFGAADAPAGAT--AGIHVSYQLYQITWLLDSEALHP 454
DB 375 ----WDGPRSHDOTLQEGVHWYFSPFRYIGISLVEPPLLHP 410

RESULT 5
O17499 PRELIMINARY; PRT: 415 AA.
AC O17499;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
ID Amphihh protein.
GN AMPHIHH.
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7739;
RN [1]
RP SEQUENCE FROM N.A.
RA Shireld S.M.;
RT "Sequence and expression of an amphioxus hedgehog gene.";
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: Y13858; CAA74169.1; -.
DR HSSP: 062226; LVNH
DR InterPro: IPR001767; Hedgehog_hint.
DR InterPro: IPR003587; Hedgehog_hincn.
DR InterPro: IPR003320; HH_signal.
DR InterPro: IPR002203; Inteln.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01085; HH_signal; 1.
DR Pfam: PF01079; Hnt; 1.
DR PRINTS: PR00632; SONICHHOG.
DR Prodom: PD003042; HH_signal; 1.
DR SMART: SM00305; HincC; 1.
DR SMART: SM00306; HincN; 1.
SQ SEQUENCE 415 AA; 46765 MW; 568D2BD2158DD931 CRC64;

```



Best Local Similarity 50.0%; Pred. No. 2,7e-73;  
Matches 229; Conservative 53; Mismatches 119; Indels 57; Gaps 8;

QY 3 LLACLL-LTVYSSLLVCSG--LACGPGRGFKRRHPKLLPLAYKOFIPNVAEETIGAS 59  
Db 4 VLAAMMYLVLAISALGHGWSEACGPGRGFKRRHPKLLPLAYKOFIPNVAEETIGAS 63  
QY 60 GYAEKISRNSEFEKELTPNYNPDIIFKDEENTGADRLMTORCKDLALAIYVNMQWPG 119  
Db 64 GLFNRIITROSEFHTLKNNTDIIIFKDEETGADREMTORCKDLALAIYVNMQWEG 123  
QY 120 VKLRYTEGWDEGHHSESLHYEGRAVDITTSDRSKRYGMALRLAVAGFDWYVESKA 179  
Db 124 VKLRYTEGWDEGHHSESLHYEGRAVDITTSDRSKRYGMALRLAVAGFDWYVESKA 183  
QY 180 HIKSVKAKENSVAASGCGFPGSATVHLEOGGKTLVKDLSPGDRVLADDDGRLLYSFEL 239  
Db 184 HIKSVKAKENSOTATQGGCFSAESVWTDGDNRIKMRVPRGDKYLSMDSGHPFSEVL 243  
QY 240 TFLDRDDGAKKVFYVETREPRERLLTAHLLEFVAPHNDSATGEPASGSGPPSGAL 299  
Db 244 TFMRESRGPMVYTIHTDDRNITATPASHLVETESRDLS----- 286  
QY 300 GPR-ALFASRRRPGQRYVVAEROGDRLLPAAYHVSYTLSEAGATAPLTAQITILNR 358  
Db 287 -PRIAKFNSDARPEFELLTPDSGGFR--KVKTIVSYTMBE-KGAYAPLVHGVVDN 342  
QY 359 VLACVAVIEESHNAHFAFAPFLAHLAALAPARTDGGDSGGGGGGRVALTAP 418  
Db 343 VAMSCYALIESQALAHWFAPFLRYQLTSSL----- 374  
QY 419 GAADAPGAGAT--AGIHWSQLVQIGTWLDSALHP 454  
Db 375 --WDGPRSHDQTLQEGVHWYPSFEYRYGSLVEPTLLHP 410

## RESULT 6

QY 073803 PRELIMINARY; PRT: 442 AA.  
AC 073803;  
DT 01-AUG-1998 (Tremblrel. 07, Created)  
DT 01-AUG-1998 (Tremblrel. 07, last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)  
DE Fugu hedgehog.  
GN FHL.  
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;  
OC Tetraodontidae; Takifugu.  
OX NCBI\_TaxID=31033;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gelliner K., Brenner S.,  
RT MEDLINE=99177347; PubMed=10077531;  
RT "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu  
rubripes.";  
RL Genome Res. 9:251-258(1999).  
DR EMBL: AF056116; AAC34384.1; -.  
DR HSP: Q62226; IYHH.  
DR InterPro: IPR001767; Hedgehog\_hnt.  
DR InterPro: IPR003586; Hedgehog\_hntC.  
DR InterPro: IPR003587; Hedgehog\_hntN.  
DR InterPro: IPR000320; HH\_signal.  
DR InterPro: IPR002203; IntraIn.  
DR Pfam: PF01085; HH\_signal; 1.  
DR Pfam: PF01079; Hnt; 1.  
DR ProDom: PD003042; HH\_signal; 1.  
DR SMART: SM00305; HntC; 1.  
DR SMART: SM00306; HntN; 1.  
SQ SEQUENCE 442 AA; 49286 MW; 1CB2B423B4B748E8 CRC64;

Query Match 42.5%; Score 1049.5; DB 13; Length 442;

Best Local Similarity 48.5%; Pred. No. 1,2e-70;  
Matches 224; Conservative 67; Mismatches 132; Indels 39; Gaps 9;

QY 4 LARCLLVYSSLLVCSLACGPGRGFKRRHPKLLPLAYKOFIPNVAEETIGASGYE 63  
Db 9 LAQVSLLAWSCVMLVOC--CGPGGIGIRRRPRKLMYKOFIPNVAEETIGASGAE 66  
QY 64 GKISRNSEFEKELTPNYNPDIIFKDEENTGADRLMTORCKDLALAIYVNMQWPGVLR 123  
Db 67 GKIRNSENFEKELTCNVPDIIFKDEENTNADREMTORCKDLALAIYVNMQWGVLR 126  
QY 124 VTEGWDEGHHSESLHYEGRAVDITTSDRSKRYGMALRLAVAGFDWYVESKAHIC 183  
Db 127 VTEGWDEGHHSESLHYEGRAVDITTSDRSKRYGMALRLAVAGFDWYVESKAHIC 186  
QY 184 SVKAKENSVAASGCGFPGSATVHLEOGGKTLVKDLSPGDRVLADDDGRLLYSFEL 243  
Db 187 SVKADHSAVAEKGCGFPWSKRYVAGGQKSLSTLPDRVMAUSEIGQVFSFVLFELH 246  
QY 244 RDDGAKKVFYVETREPRERLLTAHLLEFVAPHNDSATGEPASGSGPPSGALGPA 303  
Db 247 RDPESRMFELSLQTEDGR-RLAVTPHLLVFSDAH-----CGPDSQY--QA 289  
QY 304 LFASVRPGQRYVVAEROGDRLLPAAYHVSYTLSEAGATAPLTAQITILNRVLASC 363  
Db 290 QFASRAQTGCVLV--HTAGGEVHPSRIVSIT-EESVGAVALTEAGSVFVGVLAASS 345  
QY 364 YAVIEESHNAHFAFAPFLAHLAALAPARTDGGDSG-----GGDRGGGGGGRVA 414  
Db 346 YALVEDHGLAHMAKGPVALLSSVQLMAEPENSDGSKTLPQPHALVGRD-----KVC 400  
QY 415 L--TAPGAADAPGAGATAGIHWSQLVQIGTWLDSALHP 454  
Db 401 ARNSTSVASEAGPRGRTSEVHWYQLHRLGMYLVNPLDLP 442

## RESULT 7

QY 09GRAB PRELIMINARY; PRT: 447 AA.  
AC 09GRAB;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)  
DE Hedgehog protein (Fragment).  
GN Gryllus bimaculatus (Two-spotted cricket).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae;  
OC Gryllinae; Gryllus.  
OX NCBI\_TaxID=6999;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Niwa N., Inoue Y., Nozawa A., Saito M., Mitsuiri Y., Ohuchi H.,  
RT Yoshitaka H., Noji S.,  
RT "Correlation of diversity of leg morphology in Gryllus bimaculatus  
(cricket) with divergence in dpp expression pattern during leg  
development.";  
RL Development 127:4373-4381(2000).  
DR EMBL: AB044709; BAB19658.1; -.  
DR HSP: Q62226; IYHH.  
DR InterPro: IPR001767; Hedgehog\_hnt.  
DR InterPro: IPR003586; Hedgehog\_hntC.  
DR InterPro: IPR003587; Hedgehog\_hntN.  
DR InterPro: IPR000320; HH\_signal.  
DR InterPro: IPR001657; SonicHH.  
DR Pfam: PF01085; HH\_signal; 1.  
DR Pfam: PF01079; Hnt; 1.  
DR PRINTS: PRO0632; SONICHHOG.  
DR ProDom: PD003042; HH\_signal; 1.  
DR SMART: SM00305; HntC; 1.  
DR SMART: SM00306; HntN; 1.  
FT NON\_TER 1  
SQ SEQUENCE 447 AA; 48008 MW; 29AEPB061C3BEF0 CRC64;



Query Match	42.28;	Score 1041.5;	DB 5;	Length 447;
Best Local Similarity	51.48;	Pred. NO. 5e-70;		
Matches 228;	Conservative 47;	Mismatches 118;	Indels 51;	Gaps 10

OY		1	M L L A R C I L L V Y S S L I V C S G I A C G P G G F E K R H R P K K L P L A Y K O F I P N V A E K T T G A S G	60
Db		15	M P L L I R L L I L L L --- L M O G A R A C C G P G R A C G R R A S R L T P L V F K O H P A N V S E H T T G A A G	71
OY		61	R Y E G K I S N S E R F E L P T P N Y N D P I I F K D E N T G A D R L M T O R C K D K L A L I S V A N O M P G V	120
Db		72	P A E R R V A D D P R F D L V P N Y N A D I V F K D E B E G T G A D R L M T O R C K E K L N T A I S V A N O M P G V	131
OY		121	K L A R T E G H D E G C H H S E S B L H E G R A V D I T S D R D R S K Y G M L A R L A V A G C D W Y Y E S K A H	180
Db		132	R L A R Y E G D E G G H A A D S L H E G R A V D T T S D R D R S K Y G M L A R L A V A G C D W Y Y E T R G H	191
OY		181	I H C S Y K A E N S Y A A K S G C G F P G S A I V H L E O G G T K L V K D L S P G D R - V L A A D Q G R L L Y S D F L	239
Db		192	I H C S Y K S S S Q A A S G C G F A A E S Y V O - T P G L C A L A E L R A G R R C A G A G P G H G R I A F S P V L	250
OY		240	T F L D R D C A K K V F Y V I E T R E P R E R L L T A H L L F V A P H N D S A T G E P E A S S G S G P S G G A L	299
Db		251	L F L D R D P - A P R T L L R V R T A S G R - T L A T P S H I L L ----- P A A R G G G C F -----	290
OY		300	G P R A L F A S R V P P G Q R V V Y V A A R E D D R L L P A A V H S Y T L S E A A G A V A P L A O G F I L I N R V	355
Db		291	- P E A R F A D A V P R G D A L L V A D A G A V R - - P R V L H V D A E A T R G S V A P L T A A G T V A V D G V	347
OY		360	L A S C Y A V I E H S W A H R A F A P R L A H A L L A L A P A R T D R G D S G G D R G G G G G G A L T A P G	419
Db		348	L A S C Y A A V G S H L A H M S F A P A V R A W H - M L T A W G H A A P D Y -----	388
OY		420	A A D A G A G A T A G I H Y S Q L L Q I G 443	
Db		385	A H P P P A A A C G V H W Y A K A L S L G 408	

ID	061676	PRELIMINARY:	PRT:	410 AA.
AC	061676;			
DT	01-AUG-1998 (TREMBLrel. 07, Created)			
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Hedgehog.			
OS	Lycenchinus variegatus (sea urchin).			
CC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinzoa;			
CC	Echinoidae; Euechinoidae; Echinacea; Temnopneustoidae;			
CC	Lycenchinus.			
OX	NCBI_TaxID=7654;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Hertzler P.L., McClay D.R.;			
RT	"A sea urchin hedgehog homolog.,"			
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF059606; AAC15065.1; -.			
DR	HSSP: Q62226; 1VH.			
DR	InterPro: IPRO01767; Hedgehog_hint.			
DR	InterPro: IPRO03586; Hedgehog_hinc.			
DR	InterPro: IPRO03587; Hedgehog_hintn.			
DR	InterPro: IPRO00320; HH_signal.			
DR	InterPro: IPRO02203; Inclin.			
DR	InterPro: IPRO01657; SonicHH.			
DR	Pfam: PFO1085; HH_signal; 1.			
DR	Pfam: PFO1079; Hint; 1.			
DR	PRINTS: PR00632; SONICHHOG.			
DR	PRODOM: PD003042; HH_signal; 1.			
DR	SMART: SM00305; Hintc; 1.			
DR	SMART: SM00306; Hintn; 1.			
QO	SEQUENCE 410 AA; 4601 MW; 875CF889DA6FDBED CRC64;			

Query match 37.18; Score 916; DB 5; Length 410;

Best Local Similarity 45.58; Pred. NO. 1.2e-60;  
Matches 206; Conservative 65; Mismatches 128; Indels 54; Gaps 11;

```

0Y 4 LARLLLYVSSL-IVCSLGGPGRGKRRH-PKLLPLAYKOFIPNVAEKTLGASGR 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 MMYLLVQITLACTLALTOACHPERS-GKSHRRNRNTPLOYTKQVPM:SDTFCASGP 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 62 YEGKISHNSERFKELPNYNPDIIFKDEENFGADRLMTORCKDLNALAISVNNQMPGVK 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 PEGRIDRDERDFSKLSNNNDIVFKDEEGFADRLMTORCKDKRLNTLAISVNNEMPGIK 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 122 LRYTEGDEGCHSESLYBEGRAVDITTSDBDSKTKGMLARLAVEGFPWYTESKAI 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 125 LRVAEADDEQ-NPNRPPLHAEGRAVDITTSDBDKNKKGALRLAIVEGFPWYTESKAV 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 182 HCSVKAENSVAAKSGGCFPGSATVHLEOGGTRKLVKDISPCDGYLAADQGRLLYSDFLTF 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 184 HCSVKSLSMAAKKSGGCFPGFSQATLKNKGRMISMULDIVGDEVAVVNNNELDYSDVIMI 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 242 LDRDGAKKVYVYIEFRERERILLTAHLLFVAPRHNDASGTGEASGSGPSPGALCP 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 244 VHRKLNDSTLYFIET-EKKTIVOLTPOHLLYVS-----ERESNPDQS----- 288
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 302 RALPASRVRRGCGVYVVAERDDRRLPLAAVSHVTLSEAGAVALPLAGCTILINRYLA 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 286 RAIFASAEVRLNOCVYVTTAA-NHDKGRPRPKAVSYT-TRLGHTAAAPYTRGSLVIDGVAV 344
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 362 SCYAVIEEHSMAHRAFPRLNAHLLAALAAPARDRGDGGSGGGRGGGVALLTAPGAA 422
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 344 SSVAVMDEWAVAHASAPIRWYTYISHMLGITDPTDGTGEO----- 388
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 422 DABGAGTAGIHWYSOLLQIQTWLLDSALHP 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 384 -----RVHWYTGGLKLRKYVA-SDRLFP 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 9
ID 09R179 PRELIMINARY; PRT: 161 AA.
AC 09R179;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sonic hedgehog protein (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY.
RA Garces P.L., Meyer R.A. Jr., Brown C.A., Price D.K.;
RT "Sonic hedgehog in the rat.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR SUBMIT: AF162915; AAD45373.1; -.
DR HSSP: 062226; 1VHH.
DR MEROPS: C46.002; -.
DR InterPro: IPR001767; Hedgehog_hint.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01085; HH_signal; 1.
DR Pfam: PF01079; Hint; 1.
DR PRINTS: PR00632; SONICHHOG.
DR Prodom: PD003042; HH_signal; 1.
FT NON_TER 1
FT TER 161
SQ SEQUENCE 161 AA; 17906 MW; DF06D85FD45AE4DB CRC64;

```

```
Query Match      34.4%; Score 850; DB 11; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.8e-56;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 84 IIFKDEENTGADRLMTQRCCKDKLNALASVMNQWPGVKLRVTEGWDDEGHSEESLHYEG 143

```
Db 1 IIFKEENTGADRLMTQRCCKDLNALAISVNMQGVKRLRYEGDEGHSEESLHYEG 60
OY 144 RAYDITTSDDRBSKYGMLARLAVEAGFDWVYYESKAHHCYSKAKNSVAASGGCFPSA 203
Db 61 RAYDITTSDDRBSKYGMLARLAVEAGFDWVYYESKAHHCYSKAKNSVAASGGCFPSA 120
OY 204 TVHLEOGGCTKLVKDLSPGDRVLAADQGRLLYSDFLTFELDR 244
Db 121 TVHLEOGGCTKLVKDLSPGDRVLAADQGRLLYSDFLTFELDR 161

RESULT 10
OYGV7 PRELIMINARY; PRT; 150 AA.
ID OYGV7;
AC OYGV7;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Ambystoma mexicanum (Axolotl).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Ambystomatidae;
OC Ambystoma
CX NCBI_TaxID=8296;
RN [1]
RP SEQUENCE FROM N.A.
RA Tokok M.A., Izpitua-Beimonte J.C., Gardiner D.M., Bryant S.V.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF031480; AAD18128.1; -
DR HSSP: Q62226; IVHH.
DR InterPro: IPR001167; Hedgehog_hntc.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01085; HH_signal; 1.
DR Pfam: PF01079; Hntc; 1.
DR PRINTS: PR00632; SONICHHOG.
DR ProDom: PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 150
SQ SEQUENCE 150 AA; 16599 MW; 986F65037A69A8AC CRC64;

Query Match
Best Local Similarity 29.7%; Score 733; DB 13; Length 150;
Matches 140; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 83 DIFDEENTGADRLMTQRCCKDLNALAISVNMQGVKRLRYEGDEGHSEESLHYE 142
Db 1 DIFDEENTGADRLMTQRCCKDLNALAISVNMQGVKRLRYEGDEGHSEESLHYE 60
OY 143 GRAVDITTSDDRBSKYGMLARLAVEAGFDWVYYESKAHHCYSKAKNSVAASGGCFPS 202
Db 61 GRAVDITTSDDRBSKYGMLARLAVEAGFDWVYYESKAHHCYSKAKNSVAASGGCFPS 120
OY 203 ATVHLEOGGCTKLVKDLSPGDRVLAADQGR 232
Db 121 AKVLEHGVTRPVADLRPGDRVLAADQGR 150

RESULT 11
OYXSI6 PRELIMINARY; PRT; 139 AA.
ID OYXSI6;
AC OYXSI6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
CX NCBI_TaxID=9913;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=TOOTH GERM;
RA Koyama E., Iwamoto M., Ohmori T., Kurisu K., Wu C., Ookura T.,
RA Bashir M.M., Tucker T., Pacifici M.;
RT "Development of Stratum Intermedium and its Role as a Sonic Hedgehog-
RT Signaling Structure during Odontogenesis.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF144100; AAD33926.1; -
DR HSSP: Q62226; IVHH.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHHOG.
DR ProDom: PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 139
SQ SEQUENCE 139 AA; 15961 MW; C63B0B9682046B1 CRC64;

Query Match
Best Local Similarity 29.6%; Score 731; DB 6; Length 139;
Matches 137; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 43 AYKQITPVNAEKTLAGSGRYEGKISRNSERKELTPNPNPDIIFKDEENTGADRLMTQRC 102
Db 1 AYKQITPVNAEKTLAGSGRYEGKISRNSERKELTPNPNPDIIFKDEENTGADRLMTQRC 60
OY 103 KDKLNALAISVNMQGVKRLRYEGDEGHSEESLHYEGRAVDITTSDDRBSKYGMLA 162
Db 61 KDKLNALAISVNMQGVKRLRYEGDEGHSEESLHYEGRAVDITTSDDRBSKYGMLA 120
OY 163 RLAVEAGFDWVYYESKAH 181
Db 121 RLAVEAGFDWVYYESKAH 139

RESULT 12
OYW6C1 PRELIMINARY; PRT; 138 AA.
ID OYW6C1;
AC OYW6C1;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Sonic hedgehog protein (Fragment).
GN SHH.
OS Eleutherodactylus coqui (Puerto Rican coqui).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;
OC Eleutherodactylus.
CX NCBI_TaxID=57060;
RN [1]
RP SEQUENCE FROM N.A.
RA Carl T.F., Richardson M.K., Olsson L., Schlosser G., Klymkowsky M.W.,
RA Hanken J.;
RT "Differences in vertebrate limb development revealed by studies of the
RT direct developing frog E. coqui.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF113403; AAD23436.1; -
DR HSSP: Q62226; IVHH.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHHOG.
DR ProDom: PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15751 MW; FF4156A17F4681F0 CRC64;

Query Match
Best Local Similarity 28.7%; Score 709; DB 13; Length 138;
Matches 133; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 48 IPNVAEKTLAGSGRYEGKISRNSERKELTPNPNPDIIFKDEENTGADRLMTQRCCKDLN 107
```

```

Db 1 IPVNAKTLGASRGYEGKITRNSERKELTPNYSNIIIFKDEENTATADRLMORCKDKLN 60
Qy 108 ALAISVNMNPGVKLVNTEGMDGHHSESLHYEGRAVDITTSDBRSKYGLARLAVE 167
Db 61 ALAISVNMNPGVKLVNTEGMDGHHSESLHYEGRAVDITTSDBRSKYGLARLAVE 120
Qy 168 AGFDWYVESKAHHCVS 185
Db 121 AGFDWYVESKAHHCVS 138

RESULT 13
09WV29 PRELIMINARY: PRT: 177 AA.
ID 09WV29;
AC 09WV29;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Indian hedgehog protein (Fragment).
OS Rattus norvegicus (Rat).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RA Gages P.L., Meyer R.A. Jr., Brown C.A., Price D.K.;
RT "Indian hedgehog in rat.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR HSSP: 062226; 1VHM.
DR MEROPS: C46.003; -.
DR InterPro: IPR001767; Hedgehog_hint.
DR InterPro: IPR003587; Hedgehog_hint.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01085; HH_signal; 1.
DR Pfam: PF01079; Hint; 1.
DR PRINTS: PR00632; SONICHHOG.
DR ProDom: PD003042; HH_signal; 1.
DR SMART: SM00306; HintN; 1.
FT NON_TER 1
FT SEQUENCE 177 AA; 19739 MW; CA5871626A565F65 CRC64;

Query Match 28.4%; Score 702; DB 11; Length 177;
Best Local Similarity 74.0%; Pred. No. 4.2e-45;
Matches 131; Conservative 16; Mismatches 30; Indels 0; Gaps 0;

Qy 97 LMTORCKDKLNALAISVNMNPGVKLVNTEGMDGHHSESLHYEGRAVDITTSDBRS 156
Db 1 LMTORCKDKLNALAISVNMNPGVKLVNTEGMDGHHSESLHYEGRAVDITTSDBRN 60
Qy 157 KYGLARLAVEAGFDWYVESKAHHCVSVAKNSGCGCPGSAATVHLEGGGKLVK 216
Db 61 KYGLARLAVEAGFDWYVESKAHHCVSSESAAKTGGCPAQAQVLTLETGERVALS 120
Qy 217 DLSPGDRLVLAADQGRLLYSDFLTLDRODGAARFVYIETREPRRLTLTAHLLE 273
Db 121 AVKPSGRVLAMGEGDGNFTPSDVLIFLDREPNNRLRAFOVITDPPRRLLTPAHLLE 177

RESULT 14
096699 PRELIMINARY: PRT: 185 AA.
ID 096699;
AC 096699;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Hedgehog protein (Fragment).
HH.
OC Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

```

```

OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditylsia;
OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
OC NCBI_TaxID=39708;
RN [1]
RP SEQUENCE FROM N.A.
RA Keys D.N., Lewis D.L., Selegue J.E., Pearson B.J., Goodrich L.V.,
RA Johnson R.L., Gates J., Scott M.P., Carroll S.B.;
RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
RT evolution.";
RL Science 0:0-0(1999).
DR EMBL: AF117742; AAD08931.1; -.
DR HSSP: 062226; 1VHM.
DR InterPro: IPR000320; HH_signal.
DR InterPro: IPR001657; SonicHH.
DR Pfam: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHHOG.
DR ProDom: PD003042; HH_signal; 1.
FT NON_TER 1
FT NON_TER 185
FT SEQUENCE 185 AA; 20745 MW; E9B9F4CE2FAAE662 CRC64;

Query Match 25.1%; Score 619; DB 5; Length 185;
Best Local Similarity 68.2%; Pred. No. 7.7e-39;
Matches 116; Conservative 21; Mismatches 31; Indels 2; Gaps 2;

Qy 58 ASGRYEGKISRNSERKELTPNYPNPIIFKDEENTGADRLMORCKDKLNALAISVNMN 117
Db 3 ASGPPEGKITRDEKFRDLVPNNPNPIIDFKDDGTCGADRLMORCKDKLNALAISVNMN 62
Qy 118 PGVKLVNTEGMDGHHSESLHYEGRAVDITTSDBRSKYGLARLAVEAGFDWYVES 177
Db 63 PGVKLVNTEGMDGHHSESLHYEGRAVDITTSDBRSKYGLARLAVEAGFDWYFEN 122
Qy 178 KAHHCVSVAKNSGCGCPGSAATVHLEGGGKLVNLEGGGKLVNLA 227
Db 123 RSTIHCVSATTESSVGT-GAGCFPSGAVHTENGPXD-TASLKGNKRVLA 170

RESULT 15
080GN4 PRELIMINARY: PRT: 99 AA.
ID 080GN4;
AC 080GN4;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sonic hedgehog (Fragment).
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeoidea; Carcharhiniformes;
OC Scyllorhinidae; Scyllorhinus.
OC NCBI_TaxID=7830;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21930458; PubMed=11932743;
RA Tanaka M., Munsterberg A., Anderson W.G., Prescott A.R., Hazen N.,
RA Tickle C.;
RT "Fin development in a cartilaginous fish and the origin of vertebrate
RT limbs.";
RL Nature 416:527-531(2002).
DR EMBL: AF393835; AAM08228.1; -.
FT NON_TER 1
FT NON_TER 99
FT SEQUENCE 99 AA; 11468 MW; 831864BD834A1A20 CRC64;

Query Match 21.5%; Score 531; DB 13; Length 99;
Best Local Similarity 98.0%; Pred. No. 1.3e-32;
Matches 97; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 78 PNYPNPIIFKDEENTGADRLMORCKDKLNALAISVNMNPGVKLVNTEGMDGHHSE 137
Db 1 PNYPNPIIFKDEENTGADRLMORCKDKLNALAISVNMNPGVKLVNTEGMDGHHSE 60
Qy 138 SLHYEGRAVDITTSDBRSKYGLARLAVEAGFDWYVE 176

```

Db 61 SLHREGRAVDITSDRDRKTKGMLARLAVEAGDDMYYE 99

Search completed: February 20, 2003, 10:11:53  
Job time : 32.5535 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 10:10:28 ; Search time 8.52655 Seconds  
(without alignments)  
1423.285 Million cell updates/sec

Title: US-09-827-110A-15  
Perfect score: 2469  
Sequence: 1 MLLARCLLVSLVSLVCS.....GMAVSKSRGAGGAREGA 475

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US10\_PUBCOMB pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2467	99.9	475	8	US-08-900-220C-15
2	2467	99.9	475	9	US-09-883-848A-15
3	2467	99.9	475	10	US-09-021-660A-39
4	2467	99.9	475	10	US-09-151-999-15
5	2407	97.5	462	9	US-09-733-63A-14
6	2068	83.8	437	9	US-09-969-520A-1
7	2068	83.8	437	9	US-10-013-310-1
8	2068	83.8	437	9	US-09-733-63A-16
9	2068	83.8	437	10	US-09-021-660A-37
10	2065	83.6	437	8	US-08-900-220C-13
11	2065	83.6	437	9	US-09-883-848A-13
12	2065	83.6	437	10	US-09-151-999-13
13	2062	83.5	437	9	US-09-990-046-14
14	2062	83.5	437	9	US-09-969-520A-3
15	2059	83.4	437	9	US-09-969-520A-10
16	2059	83.4	437	9	US-09-969-520A-11
17	2058	83.4	437	9	US-09-969-520A-2
18	2054	83.2	437	9	US-09-969-520A-9
19	2052	83.1	437	9	US-09-969-520A-4

20	2048	82.9	437	9	US-09-969-520A-8	Sequence 8, Appl1
21	2045	82.8	437	9	US-09-969-520A-6	Sequence 6, Appl1
22	2036	82.5	437	9	US-09-969-520A-7	Sequence 7, Appl1
23	2019	81.8	437	9	US-09-969-520A-5	Sequence 5, Appl1
24	1807	73.2	425	8	US-08-900-220C-10	Sequence 10, Appl1
25	1807	73.2	425	9	US-09-883-848A-10	Sequence 10, Appl1
26	1807	73.2	425	10	US-09-021-660A-34	Sequence 34, Appl1
27	1807	73.2	425	10	US-09-151-999-10	Sequence 10, Appl1
28	1506	61.0	418	8	US-09-021-660A-38	Sequence 38, Appl1
29	1467	59.4	418	8	US-08-900-220C-14	Sequence 14, Appl1
30	1467	59.4	418	9	US-09-883-848A-14	Sequence 14, Appl1
31	1467	59.4	418	10	US-09-151-999-14	Sequence 14, Appl1
32	1432	58.0	416	8	US-08-900-220C-18	Sequence 18, Appl1
33	1432	58.0	416	9	US-09-883-848A-18	Sequence 18, Appl1
34	1432	58.0	416	10	US-09-151-999-18	Sequence 18, Appl1
35	1271	51.5	411	8	US-08-900-220C-16	Sequence 16, Appl1
36	1271	51.5	411	9	US-09-883-848A-16	Sequence 16, Appl1
37	1271	51.5	411	10	US-09-151-999-16	Sequence 16, Appl1
38	1266	51.3	411	9	US-09-733-63A-28	Sequence 28, Appl1
39	1262	51.1	449	9	US-09-990-046-29	Sequence 29, Appl1
40	1262	51.1	449	9	US-09-733-63A-30	Sequence 30, Appl1
41	1234	50.0	411	8	US-08-900-220C-12	Sequence 12, Appl1
42	1234	50.0	411	9	US-09-883-848A-12	Sequence 12, Appl1
43	1234	50.0	411	10	US-09-151-999-12	Sequence 12, Appl1
44	1183	47.9	396	8	US-08-900-220C-11	Sequence 11, Appl1
45	1183	47.9	396	9	US-09-990-046-13	Sequence 13, Appl1

## ALIGNMENTS

RESULT 1  
US-08-900-220C-15  
Sequence 15, Application US/08900220C  
Patent No. US20020045206A1  
GENERAL INFORMATION:  
APPLICANT: Miao, Ningning  
Wang, Monica  
Mahanthappa, Nagesh K.  
Jin, Ping  
Jin, Ping  
TITLE OF INVENTION: Method of Treating Dopaminergic and GABA-nergic Disorders  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: POLY, HOAG & ELIOT LLP  
STREET: ONE POST OFFICE SQUARE  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Ascii (text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/900,220C  
FILING DATE: 24-Jul-1997  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: ONV-044, 01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-08-900-220C-15

Query Match 99.9%; Score 2467; DB 8; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1,4e-192;

Matches 475: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTILGASG 60
DB 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTILGASG 60
OY 61 RREGKISRNSERFKELTPNYNDIIFKDEENTGADRLMTQROCKDKLNALAI SVMNQPGV 120
DB 61 RREGKISRNSERFKELTPNYNDIIFKDEENTGADRLMTQROCKDKLNALAI SVMNQPGV 120
OY 121 KLRVTEGMDEDHSEESLHYEGRAVDITTSRDORSKYGMRLARLAVEAGFDVYYESKAH 180
DB 121 KLRVTEGMDEDHSEESLHYEGRAVDITTSRDORSKYGMRLARLAVEAGFDVYYESKAH 180
OY 181 HCSYKAENSVAAKSGGCGPGSATVHLBOGCTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240
DB 181 HCSYKAENSVAAKSGGCGPGSATVHLBOGCTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240
OY 241 FLDRDGAKKVYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGALG 300
DB 241 FLDRDGAKKVYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGALG 300
OY 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAOGTILINRYL 360
DB 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAOGTILINRYL 360
OY 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGGDRGGGGRVALTPARGA 420
DB 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGGDRGGGGRVALTPARGA 420
OY 421 ADAPGAGATAGIHWYSQLLYOIGTWLDESEALHPLGMAVKSXSXSGAGGAREGA 475
DB 421 ADAPGAGATAGIHWYSQLLYOIGTWLDESEALHPLGMAVKSXSXSGAGGAREGA 475

```

# RESULT 2

US-09-883-848A-15

Sequence 15, Application US/09883848A

Publication No. US20030022819A1

GENERAL INFORMATION:

APPLICANT: Ling, L.

APPLICANT: Sanicola-Nadel, M.

TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES

FILE REFERENCE: CIBT-P01-119

CURRENT APPLICATION NUMBER: US/09/883, 848A

PRIOR APPLICATION NUMBER: 2001-06-18

PRIOR FILING DATE: 60/211, 919

NUMBER OF SEQ ID NOS: 48

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 15

LENGTH: 475

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (463)

OTHER INFORMATION: Xaa-unknown amino acid residue

US-09-883-848A-15

Query Match 99.9%; Score 2467; DB 9; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1,4e-192;

Matches 475: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTILGASG 60
DB 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTILGASG 60

```

```

OY 61 RREGKISRNSERFKELTPNYNDIIFKDEENTGADRLMTQROCKDKLNALAI SVMNQPGV 120
DB 61 RREGKISRNSERFKELTPNYNDIIFKDEENTGADRLMTQROCKDKLNALAI SVMNQPGV 120
OY 121 KLRVTEGMDEDHSEESLHYEGRAVDITTSRDORSKYGMRLARLAVEAGFDVYYESKAH 180
DB 121 KLRVTEGMDEDHSEESLHYEGRAVDITTSRDORSKYGMRLARLAVEAGFDVYYESKAH 180
OY 181 HCSYKAENSVAAKSGGCGPGSATVHLBOGCTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240
DB 181 HCSYKAENSVAAKSGGCGPGSATVHLBOGCTKLVKDLSPGDRVLAADDOGRLLYSDFLT 240
OY 241 FLDRDGAKKVYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGALG 300
DB 241 FLDRDGAKKVYVETREPRERLLTAHLLFVAPHNDSATGEPEASSGSGPPSGALG 300
OY 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAOGTILINRYL 360
DB 301 PRALFASRVPRGQRYVVAERDGRRLPAAVHSVTLSEEAAGAYAPLTAOGTILINRYL 360
OY 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGGDRGGGGRVALTPARGA 420
DB 361 ASCYAVIEHSHWAHRAFAFRLAHALLAALAPARTDRGDSGGGDRGGGGRVALTPARGA 420
OY 421 ADAPGAGATAGIHWYSQLLYOIGTWLDESEALHPLGMAVKSXSXSGAGGAREGA 475
DB 421 ADAPGAGATAGIHWYSQLLYOIGTWLDESEALHPLGMAVKSXSXSGAGGAREGA 475

```

# RESULT 3

US-09-021-660A-39

Sequence 39, Application US/09021660A

Patent No. US20010041668A1

GENERAL INFORMATION:

APPLICANT: Baron, M.

APPLICANT: Belausoff, M.

TITLE OF INVENTION: METHODS FOR MODULATING HEMATOPOIESIS AND VASCULAR

TITLE OF INVENTION: GROWTH

FILE REFERENCE: HUIP-P01-060

CURRENT APPLICATION NUMBER: US/09/021, 660A

PRIOR APPLICATION NUMBER: 2001-08-27

PRIOR FILING DATE: 1997-02-10

PRIOR APPLICATION NUMBER: 60/037, 513

PRIOR FILING DATE: 1997-02-10

NUMBER OF SEQ ID NOS: 42

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 39

LENGTH: 475

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (463)

OTHER INFORMATION: Xaa-unknown amino acid

US-09-021-660A-39

Query Match 99.9%; Score 2467; DB 10; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1,4e-192;

Matches 475: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTILGASG 60
DB 1 MLLARCLLVVSSLVCSGLACGPGRGFGRRHPKLTPLAYKQFIPNVAEKTILGASG 60
OY 61 RREGKISRNSERFKELTPNYNDIIFKDEENTGADRLMTQROCKDKLNALAI SVMNQPGV 120
DB 61 RREGKISRNSERFKELTPNYNDIIFKDEENTGADRLMTQROCKDKLNALAI SVMNQPGV 120
OY 121 KLRVTEGMDEDHSEESLHYEGRAVDITTSRDORSKYGMRLARLAVEAGFDVYYESKAH 180
DB 121 KLRVTEGMDEDHSEESLHYEGRAVDITTSRDORSKYGMRLARLAVEAGFDVYYESKAH 180

```

QY	181	HCVCVKKMNSVAAKSGCGFPGSTVHLEGGGTFLVYDLSPGGRVLAADDOGLTYSDELTY	240
Db	181	HCVCVKKMNSVAAKSGCGCFPGSMTVHLEGGGTFLVYDLSPGGRVLAADDOGLTYSDELTY	240
QY	241	FLDRDGCAGKVVYVYIETREPRERLLTLTAHLFLVAPHNDSATGEPEASGSGPSPSGALG	300
Db	241	FLDRDGCAGKVVYVYIETREPRERLLTLTAHLFLVAPHNDSATGEPEASGSGPSPSGALG	300
QY	301	PRALFASRVPRGGRVYVVAERDGRLLTLAAVHSTVLSSEAGAAVAPLTAGCTILLNVYL	360
Db	301	PRALFASRVPRGGRVYVVAERDGRLLTLAAVHSTVLSSEAGAAVAPLTAGCTILLNVYL	360
QY	361	ASCVAVYIEEHSWMAHRAFAPEPRLHALLAALAPARTDRGDSGGDGGGGGCVALTARGA	420
Db	361	ASCVAVYIEEHSWMAHRAFAPEPRLHALLAALAPARTDRGDSGGDGGGGGCVALTARGA	420
QY	421	ADAGACATGTHMYSOLLTQITGTLWLLDSEALHPHGMATKSSXSGAGAGAREGA	475
Db	421	ADAGACATGTHMYSOLLTQITGTLWLLDSEALHPHGMATKSSXSGAGAGAREGA	475







Db 122 RLRTGEMDEGHSEESLHYEGRAVDITTSRDNSKYGMLARLAVEAGFDWVYESKAH 181  
QY 181 IHCSYKAENSVAAKSGGCGPSATVHLEOGGTKLKVDLSPGDRVLAADDOGRLLYSDFLT 240  
Db 182 IHCSYKAENSVAAKSGGCGPSATVHLEOGGTKLKVDLRPGDRVLAADDOGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVVYVETREPRERLLTAANHLFVAPHNDSATGEPBASSGSGPPSGALG 300  
Db 242 FLDRDEGAKKVVYVETLEPRERLLTAANHLFVAPHND-----SGPPPG 286  
QY 301 PRALFASRYRPGQRYVVAERDGRRLRPAVHSTLSEAGAAYAPLTAOGTILINRYL 360  
Db 287 PSALFASRYRPGQRYVVAERDGRRLRPAVHSTLSEAGAAYAPLTAOGTILINRYL 346  
QY 361 ASCYAVIEHSHWAHRAFAFRLAHLAALAPARTDRGDSGGGGRGVALTAPGA 420  
Db 347 ASCYAVIEHSHWAHRAFAFRLAHLAALAPARTD-----GGGGGSIIP-AAOSA 395  
QY 421 ADAPGAGATAGIHWYSQLYQIGTWLDSSEALHPLGMAYKSS 462  
Db 396 TEARGAEPTAGIHWYSQLYHIGTWLDSETHMPLGMAYKSS 437  
RESULT 11  
US-09-883-848A-13  
; Sequence 13, Application US/09883848A  
; Publication No. US20030022819A1  
; GENERAL INFORMATION:  
; APPLICANT: Ling, L.  
; APPLICANT: Sanicola-Nadel, M.  
; TITLE OF INVENTION: ANGIOGENESIS-MODULATING COMPOSITIONS AND USES  
; FILE REFERENCE: CIBT-P01-119  
; CURRENT APPLICATION NUMBER: US/09/883, 848A  
; CURRENT FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/211, 919  
; PRIOR FILING DATE: 2000-06-16  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-883-848A-13  
Query Match 83.6%; Score 2065; DB 9; Length 437;  
Best Local Similarity 87.2%; Pred. No. 5.4e-160;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MILLARCLLVVSSLLVCSGLACGPGGFGRRRHPKLTPLAYKOTIPNVAEKTGLGASG 60  
Db 2 LLLLRKCFVLVIASSLLVCPGLACGPGRGFRRRHPKLTPLAYKOTIPNVAEKTGLGASG 61  
QY 61 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLNLAISVNMOMPGV 120  
Db 62 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLNLAISVNMOMPGV 121  
QY 121 KLRTVEGMDGHNSEESLHYEGRAVDITTSRDNSKYGMLARLAVEAGFDWVYESKAH 180  
Db 122 RLRTVEGMDGHNSEESLHYEGRAVDITTSRDNSKYGMLARLAVEAGFDWVYESKAH 181  
QY 181 IHCSYKAENSVAAKSGGCGPSATVHLEOGGTKLKVDLSPGDRVLAADDOGRLLYSDFLT 240  
Db 182 IHCSYKAENSVAAKSGGCGPSATVHLEOGGTKLKVDLRPGDRVLAADDOGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVVYVETREPRERLLTAANHLFVAPHNDSATGEPBASSGSGPPSGALG 300  
Db 242 FLDRDEGAKKVVYVETLEPRERLLTAANHLFVAPHND-----SGPPPG 286  
QY 301 PRALFASRYRPGQRYVVAERDGRRLRPAVHSTLSEAGAAYAPLTAOGTILINRYL 360  
Db 287 PSALFASRYRPGQRYVVAERDGRRLRPAVHSTLSEAGAAYAPLTAOGTILINRYL 346  
QY 361 ASCYAVIEHSHWAHRAFAFRLAHLAALAPARTDRGDSGGGGRGVALTAPGA 420  
Db 347 ASCYAVIEHSHWAHRAFAFRLAHLAALAPARTD-----GGGGGSIIP-AAOSA 420

Db 347 ASCYAVIEHSHWAHRAFAFRLAHLAALAPARTD-----GGGGGSIIP-AAOSA 395  
QY 421 ADAPGAGATAGIHWYSQLYQIGTWLDSSEALHPLGMAYKSS 462  
Db 396 TEARGAEPTAGIHWYSQLYHIGTWLDSETHMPLGMAYKSS 437  
RESULT 12  
US-09-151-999-13  
; Sequence 13, Application US/09151999  
; Patent No. US20020151460A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Elizabeth  
; TITLE OF INVENTION: REGULATION OF EPITHELIAL TISSUE BY HEDGEHOG-LIKE  
; FILE REFERENCE: ONY-031.02  
; CURRENT APPLICATION NUMBER: US/09/151, 999  
; CURRENT FILING DATE: 1998-08-11  
; EARLIER APPLICATION NUMBER: 08/955, 552  
; EARLIER FILING DATE: 1997-10-20  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: murine Shh  
US-09-151-999-13  
Query Match 83.6%; Score 2065; DB 10; Length 437;  
Best Local Similarity 87.2%; Pred. No. 5.4e-160;  
Matches 403; Conservative 9; Mismatches 24; Indels 26; Gaps 3;  
QY 1 MILLARCLLVVSSLLVCSGLACGPGGFGRRRHPKLTPLAYKOTIPNVAEKTGLGASG 60  
Db 2 LLLLRKCFVLVIASSLLVCPGLACGPGRGFRRRHPKLTPLAYKOTIPNVAEKTGLGASG 61  
QY 61 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLNLAISVNMOMPGV 120  
Db 62 RYEGKISRNSERFKELTPYNDPIIFKDEENTGADRLMTORCKDKLNLAISVNMOMPGV 121  
QY 121 KLRTVEGMDGHNSEESLHYEGRAVDITTSRDNSKYGMLARLAVEAGFDWVYESKAH 180  
Db 122 RLRTVEGMDGHNSEESLHYEGRAVDITTSRDNSKYGMLARLAVEAGFDWVYESKAH 181  
QY 181 IHCSYKAENSVAAKSGGCGPSATVHLEOGGTKLKVDLSPGDRVLAADDOGRLLYSDFLT 240  
Db 182 IHCSYKAENSVAAKSGGCGPSATVHLEOGGTKLKVDLRPGDRVLAADDOGRLLYSDFLT 241  
QY 241 FLDRDDGAKKVVYVETREPRERLLTAANHLFVAPHNDSATGEPBASSGSGPPSGALG 300  
Db 242 FLDRDEGAKKVVYVETLEPRERLLTAANHLFVAPHND-----SGPPPG 286  
QY 301 PRALFASRYRPGQRYVVAERDGRRLRPAVHSTLSEAGAAYAPLTAOGTILINRYL 360  
Db 287 PSALFASRYRPGQRYVVAERDGRRLRPAVHSTLSEAGAAYAPLTAOGTILINRYL 346  
QY 361 ASCYAVIEHSHWAHRAFAFRLAHLAALAPARTDRGDSGGGGRGVALTAPGA 420  
Db 347 ASCYAVIEHSHWAHRAFAFRLAHLAALAPARTD-----GGGGGSIIP-AAOSA 395  
QY 421 ADAPGAGATAGIHWYSQLYQIGTWLDSSEALHPLGMAYKSS 462  
Db 396 TEARGAEPTAGIHWYSQLYHIGTWLDSETHMPLGMAYKSS 437  
RESULT 13  
US-09-990-046-14  
; Sequence 14, Application US/09990046  
; Patent No. US20020156245A1  
; GENERAL INFORMATION:  
; APPLICANT: de Sauvage, Frederic  
; APPLICANT: Carpenter, David A.

```

: TITLE OF INVENTION: Patched-2
: FILE REFERENCE: P1405r1
: CURRENT APPLICATION NUMBER: US/09/990,046
: CURRENT FILING DATE: 2001-11-20
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/293,505
: PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-15
: NUMBER OF SEQ ID NOS: 32
: SEQ ID NO 14
: LENGTH: 437
: TYPE: prt
: ORGANISM: Mus musculus
: US-09-990-046-14

```

Query Match	83.5%;	Score 2062;	DB 9;	Length 437;
Best Local Similarity	87.0%;	Pred. No. 9.4e-160;		
Matches 402;	Conservative 10;	Mismatches 24;	Indels 26;	Gaps 3

```

0Y 1 MLLLRCLLVVSSSLVJSGSLACGPGRGKRRHPKPLPYAKOFIPNVAEKTIGASG 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 LLLLRCLLVVILASSLVYCGIACGPGRGKRRHPKPLPYAKOFIPNVAEKTIGASG 61

0Y 61 RYEGKISRNSERPKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMNQMGV 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 RYEGKITRNSERPKELTPYNPDIIFKDEENTGADRLMTORCKDLNALAISVMNQMGV 121

0Y 121 KLRVTEGMEDDGHSEESLHYEGRAVDITSDDRSKSYGMILARLAVEGPMWYVESRAH 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 KLRVTEGMEDDGHSEESLHYEGRAVDITSDDRSKSYGMILARLAVEGPMWYVESRAH 181

0Y 181 IHCYSKAENSVAASGSGCFPGSATVHLEOGGTYKLVKLSFGDRVLADDDGRLLYSDFLT 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 IHCYSKAENSVAASGSGCFPGSATVHLEOGGTYKLVKLRGDRVLADDDGRLLYSDFLT 241

0Y 241 FLDRDDGAKKVPVYVETREPRERLLTAAHLTVYAPHNDSATGEPEASSGSGCPGSGCALG 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 242 FLDRDDGAKKVPVYVETREPRERLLTAAHLTVYAPHNDSATGEPEASSGSGCPGSGCALG 300

0Y 301 PRALFASRVPRPGORYVVAERDDRLPLPAVHSVTSEEAAGAYAPLTAOGTILINVL 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 287 PSALFASRVPRPGORYVVAERDGRLLPLPAVHSVTLEEEAGAYAPLTAOGTILINVL 346、

0Y 361 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTDRGDSGGGDRGGGGRVALTARGA 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 347 ASCYAVIEEHSWAHRAFAFRLAHALLAALAPARTD-----GGGGGSIIP-AAQSA 395

0Y 421 ADAPGAGATGTHWYSOLLQYIGTWLLDSEALHPLGMAVSS 462

Db 396 TEARGAEPTAGIHWYSOLLQYIGTWLLDSETHPLGMAVSS 437

```

```

: RESULT 14
: US-09-969-520A-3
: Sequence 3, Application US/09969520A
: Patent No. US20020177163A1
: GENERAL INFORMATION:
: APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
: APPLICANT: BEACHY, Philip A.
: TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHED
: FILE REFERENCE: JHU1670-1
: CURRENT APPLICATION NUMBER: US/09/969,520A
: CURRENT FILING DATE: 2002-06-04
: PRIOR APPLICATION NUMBER: US 60/235,153
: PRIOR FILING DATE: 2000-09-22
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 3
: LENGTH: 437
: TYPE: PRT
: ORGANISM: Artificial sequence
: FEATURE:
: OTHER INFORMATION: Altered sonic hedgehog protein
: US-09-969-520A-3

```

Query Match	83.5%;	Score 2062;	DB 9;	Length 437;
Best Local Similarity	87.2%;	Pred. No. 9.4e-160;		
Matches 403; Conservative	8;	Mismatches 25;	Indels 26;	Gaps 3;

```
QY      1 MLLARCLLYLVSSLLVCSGLACGPRGGKRRHPKLTPLAYKOPITPVAKERTIGASC 60
        :|::|||::|
Db       2 LLLLARCFVLVASSLLVCBGLACGPRGGKRRHPKLTPLAYKOPTPVAKERTIGASC 61
        :|::|||::|
QY      61 RYECKISNSERFELTPNYPNDIIFKDEENTGADRLMTORCKDKLALASVMNMOPGV 120
        :|::|||::|
Db       62 RYECKITNSERFELTPNYPNDIIFKDEENTGADRLMTRCKDKLALASVMNMOPGV 121
        :|::|||::|
QY      121 KLRTTEGDDEGHSEESIHEGRNAVITTSDBDRSKYGMALARAVAGDWMYYESKAH 180
        :|::|||::|
Db       122 KLRTEGEDEGHSEESLHIEGRAVDITTSDBDRSKYGMALARAVEAGDWMYYASKAH 191
        :|::|||::|
QY      181 IHCVKAKENSYAAKSGCGFPESATVHLEOGGTGYVKDLSPGBRULAADDGRLYSDFLT 240
        :|::|||::|
Db       182 IHCVKAKENSVAASGGCGFPESATVHLEOGGTGYVKDLRGRDRIAADDGRLIYDFLT 241
        :|::|||::|
QY      241 FLDPDDGAKKYFYIETREPERELLTAHLLEYAHPNDSATGEPEASSGSGPPSGCALG 300
        :|::|||::|
Db       242 FLDPDEGAKKYFYIETLEPPERILLTAHLLEYAHPND-----SGTTPC 286
        :|::|||::|
QY      301 PRAFEASVRGQRYVYAEBEDGRRLLPAAVHSVTSEEAGAAYAPLTAOGTILINRYL 360
        :|::|||::|
Db       287 PSALFASVRGQRVYVAEKGRRLLPAAVHSVTLREEAGYAPLTHAGHTILINRYL 346
        :|::|||::|
QY      361 ASYAAVIEESHMAHRAEPFLHALLAALAPARTDRGGDGGGDCGGGGGGAALAPGA 420
        :|::|||::|
Db       347 ASYAAVIEESHMAHRAEPFLHALLAALAPARTD-----GGGGGSIP-AAQSA 395
        :|::|||::|
QY      421 ADAPGAGTAGIHMYSQLTYOIGTWLLDSLEALPHLGAAVSS 462
        :|::|||::|
Db       396 TEARKAEPTAGIHMYSQLTYIGTWLLDSETMRHLAGAAVSS 437
```

```

: RESULT 15
: US-09-969-520A-10
: Sequence 10, Application US/09969520A
: Patent No. US20020177163A1
: GENERAL INFORMATION:
: APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
: APPLICANT: BEACHY, Philip A.
: TITLE OF INVENTION: METHOD OF USE OF SONIC HEDGEHOG PROTEIN AS A LIGAND FOR PATCHE
: FILE REFERENCE: JRU1670-1
: CURRENT APPLICATION NUMBER: US/09/969,520A
: CURRENT FILING DATE: 2002-06-04
: PRIOR APPLICATION NUMBER: US 60/235,153
: PRIOR FILING DATE: 2000-09-22
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 10
: LENGTH: 437
: TYPE: PRT
: ORGANISM: Artificial sequence
: FEATURE:
: OTHER INFORMATION: Altered sonic hedgehog protein
: US-09-969-520A-10

```

Query Match	Similarity	83.4%	Score 2059	DB 9	Length 437
Best Local	Similarity	87.0%	Pred. No. 1.6e-159.		
Matches	Conservative	9	Mismatches	25	Indels
				26	Gaps
					3
Qy	1	MLLLARCLLVVSSLLVCSGLACGPGNRGFRGRHPRKLTPLATYKQFINVAEKTIGAGG	60		
		: :			
Db	2	LLLLARCLVLIASSLLVCSGLACGPGNRGFRGRHPRKLTPLATYKQFINVAEKTIGAGG	61		
Qy	61	RYEGKISRNSERFEKELTPYNPNDIIFKDEENTGADRLMTORCKDKNALATISVMNQPCV	120		
		: :			
Db	62	RYEGKITSRNSERFEKELTPYNPNDIIFKDEENTGADRLMTORCKDKNALATISVMNQPCV	121		
Qy	121	KLRVTEGMDCHHSBESELTHTGKRAVDITTSRDSKTKGLARLAVEAGFDWYVESKAH	180		

Qy	1	MLLALRCLLVLVSSLLVSCGLACGPGRGKRRHPKPLPLAVKQGIPIVAEETLIGASC	60
		:::	
Db	2	LLLLARCLVLVASSLLVCPGLACGPGRGKRRHPKPLPLAVKQGIPIVAEETLIGASC	61
Qy	61	RYEKISINSRREKELPNVNPDIIFDEENTGADRLMTQRCRDKLNALISVNNQPCV	120
		:::	
Dd	62	RIEGKILINSRREKELPNVNPDIIFDEENTGADRLMTQRCRDKLNALISVNNQPCV	121
Qy	121	KLRYTEGDEEDGHHSESLAYEGRAVDITTSDDRKSRYGMALRLAVEAGFDWVYESKAH	180

```
Db 122 KLRTEGWEDEGHSEESLHTEGRAVDITTSADRAKYGMLARLAVEAGFDMVYYESKAH 181
Oy 181 IHCSYKAENSVAKSGGCFPGSATVHLEQGGTKLVKDLSPGDRVLAADDQGRLYSDFLT 240
Db 182 IHCSYKAENSVAKSGGCFPGSATVHLEQGGTKLVKDLRPGDRVLAADDQGRLYSDFLT 241
Oy 241 FLDRDDGAKKVFYIETREPRERLLTAHLLFVA PHNDSATGEPEASSGSGPPSGALG 300
Db 242 FLDRDGAKKVFYIETLEPRERLLTAHLLFVAPHND-----SGPTPG 286
Oy 301 PRALFASRVPRGQRYVVAERDGRRLPRAVHSTLSEEAAGAYAPLTAOCTIILNRYL 360
Db 287 PSALFASRVPRGQRYVVAERDGRRLPRAVHSTLSEEAAGAYAPLTAOCTIILNRYL 346
Oy 361 ASCYAVIEHSHVAHRAFAFRLAHALLAALAPARTDRCGDSGGGGRVALTAPCA 420
Db 347 ASCYAVIEHSHVAHRAFAFRLAHALLAALAPARTD-----GGGGGSIP-AAOSA 395
Oy 421 ADAPGAGATAGIHWYSQLLYQIGTWLDESEALHPLGMAVKSS 462
Db 396 TEARGAEPTAGIHWYSQLLYHIGTWLDESETHPLGMAVKSS 437
```

Search completed: February 20, 2003, 10:14:43  
Job time : 9.52655 secs